

Università degli Studi di Padova  
Dipartimento di Scienze Statistiche  
Corso di Laurea Magistrale in  
Scienze Statistiche



**Modelli HAR per la misura della persistenza nella  
volatilità**

Relatore Prof. Luisa Bisaglia  
Dipartimento di Scienze Statistiche

Laureando: Davide Masiero  
Matricola N. 1058420

Anno Accademico 2013/2014



# Indice

<b>Introduzione</b>	<b>3</b>
<b>1. La memoria lunga</b>	<b>5</b>
1.1. Generalità	5
1.2. Motivazioni	7
1.3. La memoria lunga e la volatilità	10
1.4. Il modello ARFIMA	12
<b>2. Il modello HAR</b>	<b>15</b>
2.1. Il principio del modello	15
2.2. Il modello HAR(3)	17
2.3. Il modello HAR( $\infty$ )	19
2.4. La teoria asintotica	23
2.5. Errore di previsione del modello HAR( $p$ )	25
2.6. La funzione di autocorrelazione	27
<b>3. Esperimento Monte Carlo</b>	<b>41</b>
3.1. Obiettivi	41
3.2. Procedura di analisi	43
3.3. I risultati	46

<b>4. Applicazione su dati reali</b>	<b>55</b>
4.1. La serie storica .....	55
4.2. Confronto tra i modelli .....	59
<b>Conclusione</b>	<b>69</b>
<b>Bibliografia</b>	<b>71</b>

# Introduzione

Il fenomeno della memoria lunga, o dipendenza di lungo periodo, era già noto alla fine del 1800, ma solo recentemente in letteratura si è studiato in maniera più approfondita.

Si è osservato che questa tipologia di dipendenza si può riscontrare nella realtà in svariati campi come, ad esempio, l'astronomia (Newcomb (1886), Jeffrey (1939)), la chimica (Student (1927)), l'agronomia (Smith (1938), Whittle (1956, 1962)), la fisica (Taqqu (1979), Mandelbrot (1975)), la meteorologia (Haslett e Raftery(1989)), la sismologia (Ogata e Abe (1991)). In particolare, è rilevante evidenziare che, anche in ambito finanziario, spesso si osservano serie storiche con questa proprietà, soprattutto nei dati ad alta frequenza. Perciò, grazie anche ai moderni avanzamenti tecnologici, che recentemente hanno permesso di trattare dati con frequenze infragiornaliere elevate, è accresciuta sempre più negli ultimi anni la necessità di riuscire a modellare nel migliore dei modi tale persistenza.

Questa tesi si concentra proprio in ambito finanziario, trattando nello specifico serie storiche relative alla volatilità. Quest'ultima è una variabile di grande interesse, essa infatti può essere interpretata come una misura del rischio e quindi è strettamente legata alle decisioni in ambito di asset allocation, risk management, option pricing. Diventa allora di fondamentale importanza trovare procedure che riescano a stimarla e prevederla con più precisione possibile, riuscendo così a sfruttare al meglio l'elevato contenuto informativo derivante da una lunga memoria.

A tal proposito è stato proposto il modello HAR(3), *Heterogeneous AutoRegressive model*, sviluppato da Corsi (2009) successivamente generalizzato da Hwang e Shin (2014), che hanno introdotto il modello eterogeneo autoregressivo di ordine infinito, HAR( $\infty$ ).

Lo scopo di questa tesi è quello di confrontare i due suddetti modelli, cercando di capire se effettivamente considerando un modello HAR di ordine infinito (o comunque con valori di  $p > 3$ ) si riesca a cogliere maggiormente la persistenza della volatilità, ottenendo così risultati migliori sia in campo di stima che in quello previsivo.

La tesi è strutturata in quattro capitoli. Nel primo viene descritta la proprietà di memoria lunga, la volatilità e i modelli ARFIMA( $p,d,q$ ), che sono adatti a cogliere la lunga memoria e noti in letteratura. Questi ultimi, per tali ragioni, sono stati presi come riferimento al fine di valutare i risultati che si sono ottenuti attraverso i modelli HAR.

Nel secondo capitolo viene presentato il modello HAR(3) proposto da Corsi (2009), il modello HAR( $\infty$ ) proposto in Hwang e Shin (2014) e la teoria asintotica dello stimatore ottenuto con il metodo dei minimi quadrati dei modelli HAR. Inoltre, ci si è concentrati sull'errore di previsione e si è analizzata la funzione di autocorrelazione di tali modelli.

Nel terzo capitolo, mediante un esperimento di Monte Carlo vengono confrontati, in ambito previsivo, i modelli HAR(3), HAR(4), HAR(5), HAR(8), considerando il caso di dati generati da un processo HAR(8). Successivamente tale analisi viene ripetuta con dati generati da un processo FI( $d$ ).

Infine nel quarto capitolo si presenta un'applicazione su dati reali, in cui si confrontano, sia in campo di stima che in quello previsivo, i vari risultati ottenuti attraverso il modello ARFIMA e modelli HAR di ordine diverso.

# Capitolo 1

## La memoria lunga

### 1.1 Generalità

La modernizzazione degli strumenti informatici è spesso venuta a coincidere con importanti innovazioni nel campo della finanza. Ad esempio, a partire dall'inizio degli anni '90, la crescente velocità di elaborazione dei processori ha reso possibile sviluppare tecniche di analisi che risultavano in precedenza impensabili, come ad esempio, si sono potute sviluppare tecniche di prezzaggio per titoli derivati di tipo non standard basate sulla simulazione del prezzo del titolo sottostante.

Una delle ultime conseguenze della rapida espansione della tecnologia hardware è rappresentata dalla possibilità di archiviare e gestire database di dimensioni imponenti, grazie alla costante crescita di capienza dei supporti magnetici e all'introduzione su larga scala (e a costi contenuti) degli strumenti ottici di lettura/scrittura.

In ambito finanziario, ciò rende possibile studiare l'andamento dei mercati in una prospettiva infragiornaliera, ovvero registrare tutte le transazioni, eseguite o solo richieste, che si rilevano sui terminali degli operatori. In

letteratura, spesso si indicano tali rilevazioni come dati tick-by-tick o trade-by-trade.

Tale capacità è importante perché permette di reperire campioni di dati con un elevatissimo contenuto informativo.

Allo stesso tempo quindi, è nata anche la necessità di riuscire a sviluppare nuovi modelli statistici capaci di cogliere e replicare le principali caratteristiche intrinseche a tale tipologia di osservazioni.

Una proprietà spesso presente in questi dati e su cui ci si concentra in questo lavoro è la memoria lunga.

Tale caratteristica è conosciuta già dalla fine del XIX secolo, anche se solo recentemente statistici e probablisti hanno concentrato l'attenzione proprio su tale aspetto.

Un processo a memoria lunga è caratterizzato da una correlazione tra le osservazioni che converge a zero ad un tasso più lento di quello atteso nel caso di dati generati da un processo ARMA o da un processo Markoviano.

Questo argomento ha trovato un rapidissimo sviluppo e interesse proprio perché tale caratteristica è presente, con un'elevata diversità applicativa, in svariati campi come la statistica, l'economia, la finanza, l'agronomia, l'astronomia, la chimica, la fisica, la geologia, l'idrologia, le scienze ambientali e matematiche e l'ingegneria.

Nel seguente lavoro si svilupperà tale tematica in campo strettamente finanziario, nello specifico, si descriveranno e confronteranno modelli adatti a cogliere e riflettere questa particolare dipendenza che si presenta nelle serie storica della volatilità.



## 1.2 Le motivazioni

Se si conducono analisi inferenziali, supponendo di disporre di un insieme di osservazioni  $x_1, x_2, \dots, x_n$  identicamente distribuite, campionate agli istanti temporali  $t = 1, 2, \dots, n$ , le assunzioni che usualmente si considerano valide sono:

1. la media della popolazione,  $\mu = E[X_t]$ , esiste ed è finita;
2. la varianza della popolazione,  $\sigma^2 = Var[X_t]$ , esiste ed è finita;
3. le variabili  $X_1, X_2, \dots, X_n$  sono incorrelate fra loro, ovvero

$$\rho(i, j) = 0 \quad \forall i \neq j,$$

dove

$$\rho(i, j) = \frac{Cov(X_i, X_j)}{\sigma^2} = \frac{E[(X_i - \mu)(X_j - \mu)]}{\sigma^2}.$$

Inoltre se valgono le assunzioni 1, 2, 3 e definendo  $\hat{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  come la media campionaria si ha che:

$$Var(\hat{X}_n) = \frac{\sigma^2}{n}. \quad (1)$$

Però nel caso in cui le osservazioni siano generate da un processo a lunga memoria e quindi l'assunzione 3 non sia valida, risultati come questo non possono essere ritenuti corretti.

Tale situazione è più che plausibile, perché spesso l'assunzione di indipendenza è solo un'approssimazione della reale struttura di dipendenza esistente fra le osservazioni.

In tal caso la (1), ovvero la varianza di  $\hat{X}_n$ , assume la seguente espressione:

$$\text{Var}(\hat{X}_n) = \frac{1}{n^2} \sum_{i,j=1}^n \gamma(i,j) = \frac{\sigma^2}{n^2} \sum_{i,j=1}^n \rho(i,j) = \frac{\sigma^2}{n^2} \left[ 1 + \frac{1}{n} \sum_{i \neq j} \rho(i,j) \right]. \quad (2)$$

L'equazione (2) vale nel caso in cui vi sia memoria di breve periodo.

Tali processi possono essere adeguatamente modellati dai ben conosciuti e ampiamente utilizzati modelli ARMA o da processi di Markov, le cui correlazioni decrescono verso zero in modo esponenziale.

Un processo a memoria lunga, invece, è un processo stazionario le cui correlazioni tendono a zero seguendo un andamento iperbolico, quindi più lentamente rispetto ai modelli di breve memoria.

La varianza campionaria quindi, differisce sia dalla (1) che dalla (2), infatti, essendo caratterizzata da una dipendenza di lungo periodo, essa converge a zero più lentamente di  $n^{-1}$ .

Asintoticamente si ha che:

$$\text{Var}(\hat{X}_n) \sim C_{var} n^\alpha, \quad \text{quando } n \rightarrow \infty,$$

dove  $C_{var}$  è una costante positiva e  $\alpha \in (0,1)$ .

Ciò comporta che la dipendenza fra osservazioni più lontane diminuisce molto lentamente all'aumentare del ritardo  $k$ .

Tali processi sono caratterizzati dalla seguente funzione di autocorrelazione:

$$\rho(k) \sim c_\rho |k|^{-\alpha}, \quad k \rightarrow \infty, \quad (3)$$

dove  $c_\rho$  è una costante finita e positiva,  $\alpha \in (0,1)$ .

L'equazione (3) descrive solamente il comportamento asintotico delle correlazioni, non specificando il comportamento di  $\rho(k)$  ai singoli ritardi. Ciò rende più difficile l'individuazione della dipendenza di lungo periodo.

Infatti, per capire se effettivamente tale processo è caratterizzato da una memoria lunga, non è sufficiente analizzare solo pochi ritardi, ma è necessario valutare l'andamento complessivo con cui le correlazioni tendono a zero.

Anche se singolarmente i valori delle autocorrelazioni possono non essere così tanto elevati, l'effetto combinato di questi ultimi può essere rilevante nelle conclusioni inferenziali a cui si giunge.

Tale aspetto è un primo motivo per cui nasce l'esigenza di studiare particolari procedure e proporre modelli adatti a trattare questa specifica tipologia di dati.

Un altro motivo rilevante è conseguenza del fatto che trattare dati caratterizzati da una memoria lunga significa avere dati con un maggiore contenuto informativo. Infatti, maggiore è la dipendenza della variabile di interesse con il suo passato tanto più precise e affidabili saranno le eventuali previsioni che si faranno, posto ovviamente l'uso di modelli adatti a modellare tale dipendenza.

Per spiegare ulteriormente come sia fondamentale riuscire a cogliere la dipendenza di lungo periodo si può ipotizzare di modellare, erroneamente, tali processi con i noti modelli ARMA, che sono capaci di replicare un'elevata, ma corta, dipendenza con le osservazioni precedenti.

In tal caso, quindi, si avrebbe la necessità di usare un elevatissimo numero di parametri, giungendo così a un aumento dell'incertezza dell'inferenza statistica e a una più difficile interpretazione dei parametri. In campo

previsivo, inoltre, si verificherebbe una convergenza troppo veloce delle previsioni alla media del processo.

## 1.3 La memoria lunga e la volatilità

In questo lavoro si tratterà nello specifico la memoria lunga presente nelle serie storiche della volatilità, ovvero il fatto che la volatilità passata abbia un impatto persistente in quella futura.

La volatilità in ambito finanziario è una variabile di grande interesse. Essa innanzitutto può essere interpretata come una misura del rischio e quindi è fondamentale per determinare decisioni in ambito di *asset allocation*, *risk management*, *option pricing*. Diventa allora necessario trovare procedure che riescano a stimarla e prevederla.

Tale questione è di grande rilievo nel caso specifico in cui si trattino dati ad alta frequenza, che spesso generano una volatilità caratterizzata da una dipendenza di lungo periodo.

Ragionando nel tempo continuo, supponiamo che il logaritmo del prezzo di una data attività finanziaria sia descritto dalla seguente equazione differenziale stocastica:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) , \quad (4)$$

dove  $\mu(t)$  rappresenta il *drift*,  $W(t)$  è un moto Browniano,  $\sigma(t)$  indica la volatilità istantanea ed è un processo stocastico indipendente da  $W(t)$ .

Per tale processo la varianza integrata associata al giorno  $t$  corrisponde all'integrale della varianza istantanea calcolato nell'intero giorno preso in considerazione ovvero nell'intervallo  $[ t-1d , t ]$  :

$$IV^{(d)} = \int_{t-1d}^t \sigma^2(\omega) d\omega . \quad (5)$$

Dai risultati ottenuti da Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001), and Barndorff-Nielsen and Shephard (2002a, 2002b), supponendo di disporre di  $M$  prezzi infragiornalieri, si dimostra che la varianza integrata (giornaliera) si può approssimare con la somma dei quadrati dei log-rendimenti infragiornalieri.

$$RV^{(d)} = \sum_{j=1}^{M-1} r_{j,t}^2 , \quad (6)$$

$p_{j,t}$  = prezzo *jesimo* del giorno  $t$

$r_{j,t} = \log(p_{j,t}) - \log(p_{j-1,t})$  con  $j = 1, \dots, M$ .

Si precisa che, coerentemente con la notazione usata in Corsi (2009), d'ora in poi si definirà rispettivamente la volatilità integrata e la volatilità realizzata come:

$$\sigma_t^{(d)} = (IV_t^{(d)})^{1/2}, \quad (7)$$

$$RV_t^{(d)} = (\sum_{j=1}^{M-1} r_{j,t}^2)^{1/2} . \quad (8)$$

In seguito, si farà riferimento a volatilità realizzate e integrate associate a diversi orizzonti temporali più lunghi del singolo giorno, come ad esempio, una settimana, un mese, un trimestre, etc.

Tali quantità sono definite come la somma delle singole volatilità giornaliere che compongono il periodo in considerazione e successivamente

le si normalizza, dividendole per il numero di giorni che compone l'intervallo di tempo considerato.

Ad esempio la volatilità realizzata associata ad una settimana sarà pari a:

$$RV_t^{(\omega)} = \frac{1}{5} (RV_t^{(d)} + RV_{t-1d}^{(d)} + RV_{t-2d}^{(d)} + RV_{t-3d}^{(d)} + RV_{t-4d}^{(d)}) , \quad (9)$$

dove  $RV_t^{(d)}$  corrisponde alla volatilità giornaliera realizzata nel giorno  $t$ .

Nel prossimo capitolo si descriverà il modello presentato in Corsi (2009), chiamato HAR(3), *Heterogeneous AutoRegressive Model*, e una sua generalizzazione, entrambi proposti per modellare e prevedere la volatilità.

Prima però, come riferimento, si presenterà anche il noto modello ARFIMA poiché è stato ampiamente studiato in letteratura e possiede un'ottima capacità di modellare e prevedere serie storiche caratterizzate da memoria lunga, come nel caso specifico della volatilità.

## 1.4 Il modello ARFIMA

I modelli ARIMA( $p, d, q$ ) Frazionari chiamati generalmente ARFIMA( $p, d, q$ ) o FARIMA( $p, d, q$ ) sono stati introdotti indipendentemente da Hosking, e Granger e Joyeux negli anni '80.

La loro specificazione è una particolare generalizzazione dei modelli ARMA, quindi per poterli descrivere si riprendono brevemente le caratteristiche di questi ultimi.

Un processo ARMA soddisfa la seguente equazione:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} ,$$

dove  $\epsilon_t \sim w.n. (0, \sigma^2)$ , si precisa che si considera  $E[\epsilon_t] = 0$  per semplicità.

La precedente equazione si può scrivere nella seguente forma più compatta:

$$\phi(B)X_t = \theta(B)\epsilon_t, \quad \text{con } t = 1, \dots, n,$$

dove  $B$  è l'operatore ritardo.

I modelli ARMA sono noti per essere adatti a modellare serie storiche stazionarie caratterizzate da una breve memoria. Nella realtà però, spesso si incorre in serie storiche non stazionarie in media perciò, per modellarle, è doveroso applicare una trasformazione, al fine di renderle stazionarie e successivamente applicare un modello ARMA.

Appartengono a questa categoria i processi ARIMA( $p, d, q$ )

Questi ultimi soddisfano la seguente equazione:

$$\phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t, \quad \text{con } d \in N,$$

dove  $\epsilon_t \sim w.n. (0, \sigma^2)$ .

In tale situazione per ottenere un processo stazionario la serie,  $X_t$ , deve essere differenziata  $d$  volte.

Un'ulteriore generalizzazione sono i processi ARFIMA( $p, d, q$ ) che differiscono dai modelli ARIMA per il fatto che il parametro  $d \in R$ , la cui specificazione è la seguente:

$$\phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t, \quad \text{con } d \in R,$$

dove  $\epsilon_t \sim w.n. (0, \sigma^2)$ .

Il valore del parametro  $d$  è rilevante, infatti al suo variare si individuano i seguenti casi:

- $d \in (0, 0.5)$  il processo ARFIMA( $p, d, q$ ) è un processo stazionario con memoria lunga,
- $d = 0$  si ottiene l'usuale modello ARMA( $p, q$ ),
- $d \in (-0.5, 0)$  il processo ARFIMA( $p, d, q$ ) è un processo stazionario con dipendenza negativa,
- $d < 0.5$  si può dimostrare che il processo è stazionario (Hosking, 1981),
- $d > 0.5$  si ottiene il processo ARFIMA ( $p, d, q$ ) con  $d \in (-0.5, 0.5)$  dopo aver opportunamente differenziato la serie.

Nel caso particolare che  $p = q = 0$  si ottiene un processo  $X_t$ , definito rumore integrato frazionario, denotato ARFIMA(0,  $d$ , 0) o FI( $d$ ).

Come dimostrato per esempio, in Hwang e Shin (2014), la funzione di autocorrelazione di tali processi può essere approssimata dalla seguente espressione:

$$\rho(k) \sim Ck^{2d-1} . \quad 10)$$

Si nota quindi che il parametro  $d$  svolge un ruolo chiave nel definire la dipendenza del processo, ovvero se:

- $d \in (0, 0.5)$  il processo possiede una funzione di autocorrelazione che tende a zero molto lentamente per questo il processo possiede una memoria lunga,
- $d = 0$  si ha una successione di variabili *i.i.d*, o memoria corta,
- $d \in (-0.5, 0)$  il processo è stazionario con memoria intermedia o dipendenza negativa.



# Capitolo 2

## Il modello HAR

### 2.1 Il principio del modello

Il principio su cui si basano i modelli HAR, descritti nelle sezioni successive, deriva dalla presenza di vari fattori eterogenei all'interno dei mercati finanziari.

Tale caratteristica è stata dimostrata e sviluppata in letteratura da Peters (1994) in "Fractal Market Hypothesis", Lux e Marchesi (1999), Alfarano e Lux (2007), LeBaron (2006).

Questi lavori evidenziano che l'eterogeneità deriva da diverse ragioni, come l'esistenza di investitori con differenti: quantità di patrimonio da investire, attitudini al rischio, orizzonti temporali di investimento, collocazioni geografiche, disponibilità e procedure diverse nel reperire le informazioni.

Focalizzandoci sulle diversità dell'orizzonte temporale di investimento, è evidente che all'interno dei mercati finanziari vi sono investitori con

differenti obiettivi e che quindi operano nel lungo, medio o breve periodo, giungendo persino a una frequenza inferiore all'ora.

Inoltre, l'orizzonte temporale dipende anche dagli scopi dell'investimento, ad esempio vi sono investitori istituzionali, come compagnie assicurative e fondi pensioni, che ovviamente operano in periodi di tempo molto lunghi.

Conseguentemente a questa eterogeneità si creano, quindi, all'interno di un mercato finanziario diverse componenti di volatilità: di breve, medio e lungo termine, causate rispettivamente da investimenti giornalieri o con frequenza anche più elevata, settimanali, di una o più mensilità.

Diventa, pertanto, fondamentale capire le relazioni che intercorrono tra le componenti di volatilità associate a orizzonti temporali differenti.

Grazie ai risultati ottenuti in Müller et al. (1997), Arneodo, et al. (1998), Gençay e Selçuk (2004) e Lynch e Zumbach (2003), si è dimostrato che la volatilità relativa ad un lungo periodo ha un'influenza maggiore sulla volatilità di breve periodo, rispetto al contrario. Per questo motivo il comportamento della volatilità si è definito asimmetrico.

Tale risultato statistico può essere spiegato a livello economico dal fatto che gli investimenti di breve periodo sono influenzati dalla volatilità di lungo periodo, perché quest'ultima è un'indicazione dei rischi e delle variazioni attese dei prezzi futuri. Al contrario chi opera nel lungo periodo non si preoccupa allo stesso modo di ciò che succede nel breve periodo.

## 2.2 Il modello HAR(3)

In questa sezione si presenta il modello proposto da Corsi (2009), chiamato HAR(3), per modellare la volatilità giornaliera realizzata.

Per poter capire la sua specificazione, si definisce la volatilità latente parziale,  $\tilde{\sigma}_t^{(\cdot)}$ , come la volatilità generata da una determinata componente di mercato.

Il modello proposto si basa sul principio del comportamento asimmetrico della volatilità, enunciato nel precedente paragrafo, secondo cui la volatilità parziale non osservata,  $\tilde{\sigma}_t^{(\cdot)}$ , è funzione della sua realizzazione al tempo precedente ( $t - 1$ ), e del valore atteso delle componenti di volatilità parziali di orizzonte temporale maggiore.

Nel modello descritto da Corsi (2009) le tre componenti considerate sono: giornaliera, settimanale e mensile definite dalla seguente specificazione:

$$\tilde{\sigma}_{t+1m}^{(m)} = c^{(m)} + \varphi^{(m)} RV_t^{(m)} + \tilde{\omega}_{t+1m}^{(m)},$$

$$\tilde{\sigma}_{t+1w}^{(w)} = c^{(w)} + \varphi^{(w)} RV_t^{(w)} + \gamma^{(w)} E_t[\tilde{\sigma}_{t+1m}^{(m)}] + \tilde{\omega}_{t+1w}^{(w)},$$

$$\tilde{\sigma}_{t+1d}^{(d)} = c^{(d)} + \varphi^{(d)} RV_t^{(d)} + \gamma^{(d)} E_t[\tilde{\sigma}_{t+1d}^{(d)}] + \tilde{\omega}_{t+1d}^{(d)},$$

dove  $RV_t^{(m)}, RV_t^{(w)}, RV_t^{(d)}$  sono rispettivamente la volatilità realizzata osservata mensile, settimanale, giornaliera descritte nella (8) e nella (9).

$\tilde{\omega}_{t+1m}^{(m)}, \tilde{\omega}_{t+1w}^{(w)}, \tilde{\omega}_{t+1d}^{(d)}$  sono le innovazioni della volatilità e sono contemporaneamente e serialmente indipendenti con media pari a zero e varianza tale da soddisfare il vincolo di positività della volatilità parziale.

Attraverso sostituzioni ricorsive si giunge a:

$$\tilde{\sigma}_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \tilde{\omega}_{t+1d}^{(d)}. \quad (11)$$

L'equazione (11) può essere vista come un modello per la volatilità a tre fattori stocastici corrispondenti alla volatilità realizzata al tempo precedente a quello considerato.

Inoltre il processo per la volatilità latente parziale può essere definito dalla seguente equazione:

$$\tilde{\sigma}_{t+1d}^{(d)} = RV_{t+1d}^{(d)} + \omega_{t+1d}^{(d)}. \quad (12)$$

Sostituendo quindi l'equazione (12) nell'equazione (11) si ottiene che:

$$RV_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \omega_{t+1d}, \quad (13)$$

con  $\omega_{t+1d} = \omega_{t+1d}^{(d)} + \tilde{\omega}_{t+1d}^{(d)}$ .

Il modello HAR(3), basato sull'equazione (13), è un modello lineare autoregressivo e, per come sono costruiti i regressori, può anche essere espresso come un modello AR(22). E' rilevante evidenziare che, come si dimostra in Corsi (2009), sebbene formalmente sia un modello a memoria corta, riesce comunque a modellare la proprietà di memoria lunga presente nei dati.

## 2.3 Il modello HAR( $\infty$ )

In questa sezione si presenta il modello HAR( $\infty$ ), un'estensione del modello HAR(3), ottenuto aggiungendo come regressori infiniti ritardi di  $RV_{t+1d}^{(d)}$ .

L'obiettivo principale è capire se il modello HAR( $\infty$ ), considerando infiniti ritardi della volatilità realizzata, apporti miglioramenti significativi in fase di stima e previsione, rispetto al più parsimonioso HAR(3).

La specificazione generale del modello è la seguente:

$$Y_t = \beta_0 + \beta_1 Y_{t,h_1} + \beta_2 Y_{t,h_2} + \dots + \epsilon_t, \quad (14)$$

dove  $\{\beta_j : j = 0, 1, 2, \dots\}$  è una successione di numeri reali che tendono a zero,  $\{h_j : j = 0, 1, 2, \dots\}$  è una successione di numeri interi positivi che crescono fino a infinito, tali che:

$$Y_{t,h_j} = \frac{1}{h_j} (Y_{t-1} + Y_{t-2} + \dots + Y_{t-h_j}), \quad (15)$$

mentre  $\epsilon_t$  è una successione *i.i.d.* di variabili casuali con media pari a zero e varianza uguale a  $\sigma^2$ .

Nel nostro caso applichiamo tale modello alla volatilità realizzata, quindi la variabile definita nella (14) con  $Y_t$  corrisponde nel nostro caso a  $RV_t^{(d)}$ , definita in precedenza nella (8). Allo stesso modo  $Y_{t,h_j}$ , definita nella (15), rappresenta la volatilità realizzata di periodo  $h_j$ , che è uguale alla media

aritmetica delle  $h_j$  realizzazioni precedenti della volatilità realizzata giornaliera, ovvero:

$$RV_{t,h_j} = \frac{1}{h_j} (RV_{t-1}^d + RV_{t-2}^d + \dots + RV_{t-h_j}^d).$$

Esplicitando i regressori del modello (14) tramite la (15) si nota che  $Y_t$  è un processo AR( $\infty$ ):

$$\begin{aligned} Y_t &= \beta_0 + \sum_{j=1}^{\infty} \frac{\beta_j}{h_j} (Y_{t-1} + Y_{t-2} + \dots + Y_{t-h_j}) + \epsilon_t \\ &= \beta_0 + \sum_{j=1}^{\infty} \alpha_j \left( \sum_{k=h_{j-1}+1}^{h_j} Y_{t-k} \right) + \epsilon_t \\ &= \beta_0 + \sum_{j=1}^{\infty} \varphi_i Y_{t-i} + \epsilon_t, \end{aligned} \tag{16}$$

con:

$$\begin{aligned} \alpha_j &= \sum_{k=j}^{\infty} \frac{\beta_k}{h_k} \text{ per } j = 1, 2, \dots \\ \varphi_1 &= \alpha_1, \quad \varphi_{h_j+r} = \alpha_{j+1} \text{ per } r = 1, 2, \dots, h_{j+1} - h_j \text{ e } j = 1, 2, \dots \\ h_0 &= 0 \end{aligned}$$

Dall'equazione (16) si possono ricavare facilmente le condizioni di stazionarietà (A1) di  $Y_t$ , che sono le seguenti:

(A1)

a) Il polinomio:

$$A(z) = 1 - \sum_{j=1}^{\infty} \alpha_j f_j(z) = 1 - \sum_{i=1}^{\infty} \varphi_i z^i$$

tende a zero per  $|z| < 1$ , dove  $f_j(z) = \sum_{k=h_{j-1}+1}^{h_j} z^k$  per  $j = 1, 2, \dots$ ,

b)

$$\sum_{j=1}^{\infty} |\beta_j| < \infty,$$

c)  $\{\epsilon_t\}$  è una successione di variabili casuali *i.i.d.* con media zero e varianza  $\sigma^2$ .

Come si dimostra in Hwang e Shin (2014), nel caso specifico di applicazioni su casi reali di serie storiche relative alla volatilità realizzata, dove è coerente che  $\beta_j > 0 \forall j \in N$ , una condizione sufficiente e necessaria per la stazionarietà è che valga A(1) a) e la condizione A(1) b) può essere semplificata con  $\sum_{j=1}^{\infty} \beta_j < 1$  poiché  $A(z)$  è una funzione monotona decrescente e  $A(0) = 1$ ,  $A(1) = \sum_{j=1}^{\infty} \beta_j > 0$ .

Inoltre assumendo che sia valida la condizione A(1) e che quindi  $Y_t$  sia stazionaria, la serie in considerazione si può rappresentare come un processo a media mobile di ordine infinito:

$$Y_t = \sum_{j=0}^{\infty} \xi_j \epsilon_{t-j}, \quad (17)$$

dove  $\xi_0 = 1$ ,  $\xi_k = \sum_{l=0}^{k-1} \xi_l \varphi_{t-l}$  per  $k = 1, 2, \dots$ .

Dopo aver ricavato le condizioni utili a garantire la stazionarietà della serie, le condizioni necessarie per ottenere la proprietà di memoria lunga sono le seguenti:

(A2)

$$\beta_j \sim c\lambda^j \quad \text{con } \lambda \text{ tale che } |\lambda| < 1 \quad \text{e } j=1,2,\dots$$

$$h_j \sim cv^j \quad \text{con } v > 1 \text{ e } j=1,2,\dots$$

Ovvero, per com'è definito  $\beta_j$  nella condizione (A2), si ha che, all'aumentare di  $j$ , la dipendenza lineare tra la variabile risposta  $Y_t$  e i regressori  $Y_{t,h_j}$  definiti nella (15) diminuisce sempre più fino a diventare trascurabile.

Invece la condizione posta su  $h_j$  comporta che  $Y_{t,h_j}$  sia la media aritmetica calcolata su un numero maggiore di ritardi di  $Y_t$  all'aumentare di  $j$ .

Dopo queste considerazioni, si giunge all'importante risultato che, in termini previsivi, la proprietà di memoria lunga, caratteristica di un processo HAR( $\infty$ ), può essere ben colta anche da un HAR( $p$ ).

Inoltre in Hwang e Shin (2014), si evince che già con  $p = 8$  si ottiene una buona approssimazione di un HAR( $\infty$ ).

A dimostrazione di quanto detto basta pensare che per i due processi HAR( $\infty$ ) e HAR( $p$ ), ipotizzando le  $\beta_j$  note, la differenza media tra le previsioni di  $Y_{n+1}$  basate sui rispettivi due modelli è equivalente a tale espressione:

$$\sum_{j=p+1}^{\infty} \beta_j \mu_Y, \text{ con } \mu_Y = E(Y_t).$$

Ma, poiché  $\beta_j$  decade esponenzialmente con l'aumentare di  $j$ , tale quantità diventa trascurabile anche con valori di  $p$  prossimi a 8.



## 2.4 La teoria asintotica

Si consideri ora il modello HAR( $\infty$ ) esplicitato nella (14) ovvero:

$$Y_t = \beta_0 + \beta_1 Y_{t,h_1} + \beta_2 Y_{t,h_2} + \dots + \epsilon_t ,$$

dove le  $h_j$  sono date, mentre  $\beta_j, j = 0, 1, 2, \dots$  sono coefficienti non noti che devono essere stimati.

Per stimare i coefficienti del modello lo si riscrive come segue:

$$Y_t = \beta_0 + \sum_{j=1}^p \beta_j Y_{t,h_j} + a_t(p) , \quad (18)$$

in cui il valore di  $p$  dipende dalla numerosità campionaria  $n$  , e tale che:

$$a_t(p) = \epsilon_t + \sum_{j=p+1}^{\infty} \beta_j Y_{t,h_j} \equiv \epsilon_t + \eta_t(p) , \quad (19)$$

con  $Y_{t,h_j}$  definita nella (15), e con  $\epsilon_t$  successione i.i.d. di variabili casuali con media pari a zero e varianza uguale a  $\sigma^2$  .

L'equazione (18) può essere espressa anche in questo modo:

$$Y_t = X_t(p)' \beta(p) + a_t(p) ,$$

dove  $X_t(p) = (1, Y_{t,h_1}, Y_{t,h_2}, \dots, Y_{t,h_p})'$  e  $\beta(p) = (\beta_0, \beta_1, \dots, \beta_p)'$ ,

quindi in termini matriciali:

$$Y = X(p)\beta(p) + a(p) , \quad (20)$$

dove  $Y = (Y_1, Y_2, \dots, Y_n)'$ ,  $X_t = (X_1(p), X_2(p), \dots, X_p(p))'$  e  
 $a_t = (a_1(p), a_2(p), \dots, a_p(p))'$ .

Dopo essere pervenuti alla specificazione (20) del modello HAR( $\infty$ ) iniziale, si possono stimare i suoi coefficienti attraverso il metodo dei minimi quadrati.

Applicando tale procedura si giunge al seguente stimatore:

$$\hat{\beta}(p) = (X(p)'X(p))^{-1} X(p)'Y = \hat{R}(p)^{-1}\hat{r}(p), \quad (21)$$

dove  $\hat{R}(p) = \frac{1}{n} \sum_{t=1}^n X_t(p) X_t(p)'$  e  $\hat{r}(p) = \frac{1}{n} \sum_{t=1}^n X_t(p) Y_t$ .

Com'è noto per poter condurre le analisi inferenziali si deve pervenire alla distribuzione dello stimatore (21).

Per giungere a questo obiettivo vengono presi in considerazione i risultati ottenuti in Hwang e Shin (2014). In tale lavoro la consistenza dello stimatore  $\hat{\beta}(p)$  è stata dimostrata nel teorema 3.3.

Assumendo valide le condizioni (A1), (A2) e con l'ausilio della seguente assunzione:

$$(A3) \quad E[\epsilon_t]^{q+\delta} < \infty \text{ per } q > 2 \text{ e } 0 < \delta < 1.$$

E' possibile dimostrare che, se  $\rho > q(q + \delta)/(2\delta)$  e  $p^{2+\epsilon} = O(n)$  per qualche  $\epsilon > 0$ ,

$$\|\hat{\beta}(p) - \beta(p)\| \xrightarrow{p} 0 \quad \text{con } n \rightarrow \infty,$$

cioè  $\hat{\beta}(p)$  è uno stimatore consistente di  $\beta$ .

Inoltre, il teorema 3.4 in Hwang e Shin (2014), dimostra che:

$$[\hat{\beta}(p) - \beta(p)] \xrightarrow{d} N(0, \sigma^2) \text{ con } n \rightarrow \infty .$$

Tale risultato è di fondamentale importanza, infatti l'intera procedura di analisi inferenziale si basa su di esso.

Va sottolineato che i due teoremi appena enunciati sono rispettivamente un'estensione dei risultati ottenuti in Ing e Wei (2003) e Kuersteiner (2005), i quali trattano modelli autoregressivi di ordine infinito, senza, tuttavia, considerare la memoria lunga.

## 2.5 Errore di previsione del modello HAR( $p$ )

Utilizzando i risultati riportati nella sezione precedente si può far vedere che la stima di previsione un passo in avanti è la seguente:

$$\begin{aligned} \hat{Y}_{n+1}(p) &= X_{n+1}(p)' \hat{\beta}(p) \\ &= X_{n+1}(p)' \hat{R}(p)^{-1} \hat{r}(p) \\ &= X_{n+1}(p)' \hat{R}(p)^{-1} \frac{1}{n} \sum_{t=1}^n X_t(p) [X_t(p)' \beta(p) + a_t(p)] \\ &= X_{n+1}(p)' \beta(p) + f_{n+1}(p) , \end{aligned}$$

dove:

$$f_{n+1}(p) = X_{n+1}(p)' \hat{R}(p)^{-1} \frac{1}{n} \sum_{t=1}^n X_t(p) a_t(p).$$

Perciò l'errore di previsione è il seguente:

$$\hat{Y}_{n+1}(p) - Y_{n+1} = f_{n+1}(p) - a_{n+1}(p) = f_{n+1}(p) - \eta_{n+1}(p) - \epsilon_{n+1}, \quad (22)$$

dove  $\eta_{n+1}(p) = \sum_{j=p+1}^{\infty} \beta_j Y_{n+1,h_j}$ .

Dalla (22) risulta evidente che l'errore di previsione è costituito da tre termini :

- $\epsilon_{n+1}$  è il termine dell'errore futuro,
- $\eta_{n+1}(p)$  è il termine dell'errore che si genera approssimando un processo HAR( $\infty$ ) con un HAR( $p$ ),
- $f_{n+1}(p)$  è il termine dell'errore di stima dei coefficienti corrispondente a  $\hat{\beta}(p) - \beta(p)$ .

Un risultato importante, per lo scopo di tale analisi, è dato dal teorema 4.1 in Hwang e Shin (2014).

Tale Teorema afferma che, considerando valide le stesse condizioni assunte nel teorema 3.3, e per  $n \rightarrow \infty$  si ha:

- $Var(\eta_{n+1}(p)) = O(\lambda^{2p})$  con  $\lambda < 1$ ,
- $Var(f_{n+1}(p)^2) = O(p/n)$ .

Ciò dimostra che, per  $p$  sufficientemente grande,  $\eta_{n+1}(p)$  ovvero, il termine dell'errore che si genera dall'approssimazione di un processo HAR( $\infty$ ) con un HAR( $p$ ), diventa trascurabile.

## 2.6 La funzione di autocorrelazione

In questa sezione si ricava la funzione di autocorrelazione dei modelli HAR( $p$ ), interpretati come approssimazione di un HAR( $\infty$ ), che viene poi confrontata con quella dello specifico modello HAR(3).

Attraverso il loro confronto si evidenziano quindi le rispettive capacità dei modelli di replicare la proprietà di lunga memoria.

A livello teorico, dal teorema 2.2 di Hwang e Shin (2014), assumendo valide le assunzioni (A1) e (A2), si evincono i seguenti due risultati:

- in riferimento alla (17), si ha che:

$$|\xi_i| \sim ck^{-\rho}$$

- inoltre, viene fornita l'espressione della funzione di autocovarianza di un modello generico HAR( $p$ ) che risulta:

$$\gamma_k = Cov(Y_t, Y_{t+k}) = \sigma^2 \sum_{i=0}^{\infty} \xi_i \xi_{i+k} \sim ck^{-\rho}$$

con  $\rho = \frac{\log v - \log |\lambda|}{\log v}$  e  $c$  pari ad una generica costante.

Quindi si nota, per come è costruito  $\rho$ , che tale funzione dipende dai due parametri  $v$  e  $\lambda$ . Inoltre dalla condizione (A2) si ha che  $\beta_j \sim c\lambda^j$  e  $h_j \sim cv^j$ .

Ne consegue che sia il valore di  $\beta_j$  che quello dei parametri  $h_j$  sono determinanti nel definire l'andamento della funzione di autocorrelazione.

Ora, per capire nello specifico come i parametri del processo influenzino la funzione di autocorrelazione, si propone uno studio di essa a livello empirico. In particolare si confrontano le funzioni di autocorrelazione ricavate da dati simulati da modelli HAR( $p$ ) al variare del valore dei parametri presenti nei modelli.

PARAMETRO  $a$ :

Nella figura 1 si confrontano le funzioni di autocorrelazione di un HAR(3) e un HAR( $p$ ) al fine di capire se, all'aumentare di  $p$ , il modello rifletta maggiormente la proprietà di memoria lunga.

Si precisa che per condurre questa analisi i dati sono stati generati da un processo HAR( $p$ ) la cui specificazione è la seguente:

$$Y_t = 0.4 + 0.36(a) Y_{t,h_1} + 0.28(a) Y_{t,h_2} + 0.28(a) Y_{t,h_3} + \sum_{j=4}^p \beta_j Y_{t,h_j} + \epsilon_t$$

dove:

$$h_1 = 1, h_2 = 5, h_3 = 22,$$

$$h_j = 22v^{j-3} \quad \text{con } j = 4, 5, \dots, p, \quad v = 2,$$

$$\sum_{j=1}^3 \beta_j = 0.92(a),$$

$$\beta_j = 0.92(1-a) \left( \frac{\lambda^{j-4}}{\sum_{j=4}^p \lambda^{j-4}} \right) \quad \text{con } j = 4, 5, \dots, p, \quad v = 2, \quad \lambda = 0.7$$

Nell'esempio si è scelto di porre  $p = 8$  per ottenere così una buona approssimazione dell'HAR( $p$ ) ad un HAR( $\infty$ ).

Inoltre, si precisa che tale specificazione dei coefficienti,  $\beta_j$ , con  $j = 1, 2, \dots, p$ , coincide con quella del modello utilizzato in Corsi (2009) nella sezione 2.4.

Per come sono definiti  $\beta_1, \dots, \beta_p$ ,  $a$  è quel parametro che condiziona il valore dei primi tre coefficienti rispetto ai successivi  $\beta_j$  con  $j = 4, \dots, p$ .

Infatti, ad esempio, se si considera il caso limite e si pone  $a = 1$  allora  $\beta_j = 0$  con  $j = 4, \dots, 8$  ciò comporta che il processo HAR(8) iniziale si semplifica in un HAR(3).

Nel caso contrario, invece, se  $a$  decresce nell'intervallo  $(0,1)$  si ha che il valore di  $\sum_{j=1}^3 \beta_j$  diminuisce mentre quello di  $\sum_{j=4}^p \beta_j$  aumenta. Questo implica che diminuendo il valore di  $a$  si generano dati sempre più prossimi ad un processo HAR(8).

I valori attribuiti ai parametri, ovvero  $\beta_1 = 0.36, \beta_2 = 0.28, \beta_3 = 0.28$ ,  $\nu = 2, \lambda = 0.7$ , con  $\epsilon_t \sim N(0,1)$  sono stati scelti perché uguali a quelli del modello proposto in Corsi (2009), dove si è dimostrato che producono dati simili alla realtà.

Si specifica che per poter simulare osservazioni da un generico processo HAR( $p$ ), si sono inizializzate le prime osservazioni necessarie per generare i dati, ponendole pari a:  $E[Y_t] = \frac{\beta_0}{(1 - \sum_{j=1}^p \beta_j)}$ . Per questo motivo, volendo creare serie storiche con numerosità pari a 3000, si sono generati 103000 dati, al fine di scartare così i primi 100000. Si evita, in questo modo di disporre durante l'analisi, di serie che dipendano dalle condizioni iniziali poste.

Inoltre, per creare funzioni di autocorrelazione che non siano dipendenti dalle specifiche realizzazioni del PGD, si sono generate 100 serie storiche e le funzioni di autocorrelazione analizzate corrispondono alla media aritmetica delle singole funzioni di autocorrelazione delle 100 serie storiche generate.

Funzione di autocorrelazione  
al variare del parametro  $a$

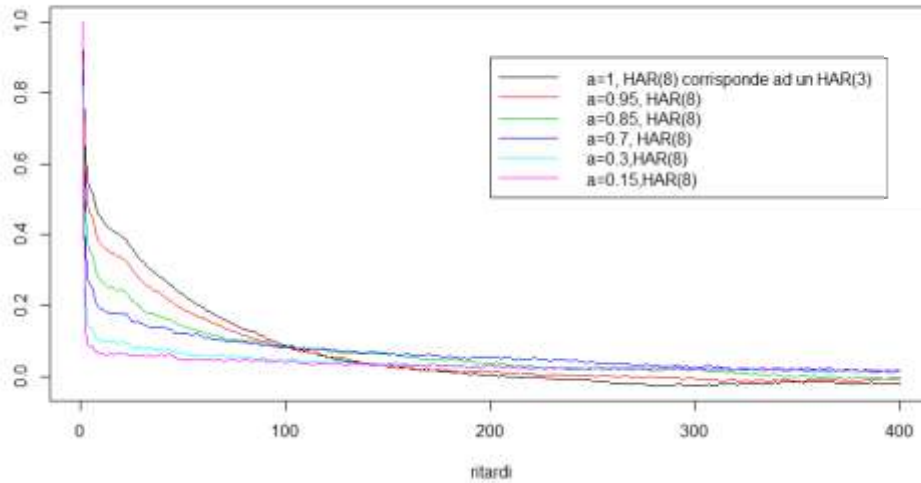


Figura 1

Dai risultati riportati in figura 1 si evince che, per tutti i valori del parametro  $a$  considerati, la funzione di autocorrelazione rimane diversa da zero per un numero elevato di ritardi (maggiore di 100). Ciò è prova del fatto che effettivamente tali modelli sono adatti a cogliere la proprietà di memoria lunga.

Osservando la figura 1 è, inoltre, evidente, che più il valore di  $a$  diminuisce, meno lentamente la funzione di autocorrelazione tende a zero.

Ciò significa che, effettivamente per un processo HAR( $p$ ) all'aumentare del suo ordine, aumenta la dipendenza di lungo periodo.



### Funzione di autocorrelazione al variare del parametro $a$

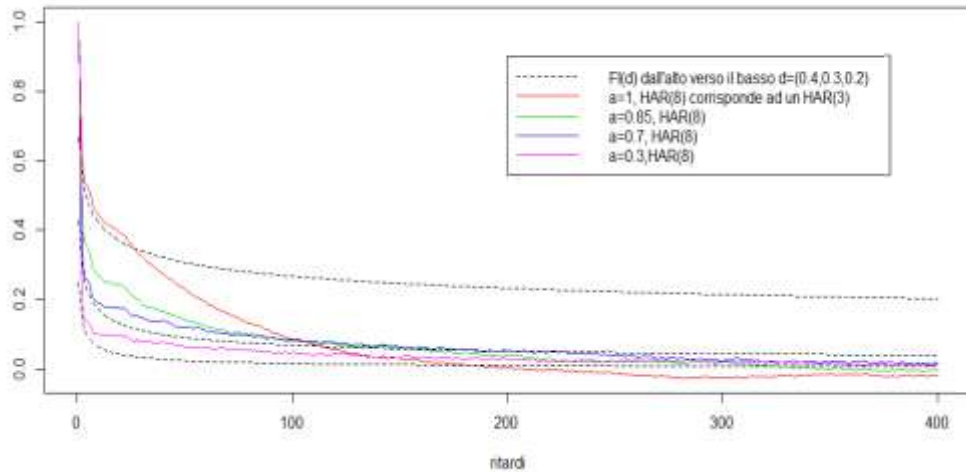


Figura 2

Nella figura 2 vengono riportate le funzioni di autocorrelazione di un processo HAR(8) con il parametro  $a$  che assume i seguenti valori: 1, 0.85, 0.7, 0.3.

Si sono riportate inoltre le funzioni di autocorrelazione generate da processi FI( $d$ ) con  $d = 0.4, 0.3, 0.2$ . Tale processo è preso come riferimento, perché adatto a rappresentare la proprietà di memoria lunga.

La figura 2 evidenzia che effettivamente un modello FI(0.4) è capace di modellare maggiormente tale proprietà rispetto agli HAR(8) considerati.

Questi ultimi, infatti, hanno funzioni di autocorrelazione più simili a un processo FI( $d$ ) con  $d$  pari a 0.3, nonostante ciò si nota che sono caratterizzati comunque da una dipendenza di lungo periodo.

Vediamo ora cosa succede facendo variare gli altri parametri presenti nel modello.

Si precisa che nelle successive analisi i dati sono generati dal modello proposto nella sezione 5.1 in Hwang e Shin (2014), specificato nel seguente modo:

$$Y_t = \beta_0 + \sum_{j=1}^p \beta_j Y_{t,h_j} + \epsilon_t$$

dove:

$$h_1 = 1, h_2 = 5, h_3 = 22,$$

$$h_j = 22v^{j-3} \quad \text{con } j = 4, 5, \dots, p,$$

$$\sum_{j=1}^p \beta_j = S_\beta,$$

$$\beta_j = \frac{S_\beta \lambda^{j-1}}{\sum_{j=1}^p \lambda^{j-1}} \quad \text{con } j = 1, 2, \dots, p,$$

$$\epsilon_t \sim N(0,1).$$

PARAMETRO  $S_\beta$  :

Il parametro  $S_\beta$  è pari a  $\sum_{j=1}^p \beta_j$ , per le condizioni di stazionarietà di un processo HAR( $p$ ) esplicitate in (A1) deve essere necessariamente  $\sum_{j=1}^p \beta_j < 1$ , inoltre, per la natura della variabile risposta che stiamo trattando, si ha che  $\beta_j > 0$  con  $j=1, \dots, p$ .

Per questi motivi  $S_\beta$  viene fatta variare nell'intervallo (0,1).

Nello specifico si sono considerati i seguenti valori di tale parametro: 0.95, 0.8, 0.6, 0.5, 0.4, 0.2 .

Si rende noto che il comportamento di  $S_\beta$  è stato valutato congiuntamente al variare del dell'ordine  $p$  dei processi HAR.

Si sono presi in considerazione i processi HAR(3), HAR(4), HAR(5), HAR(8).

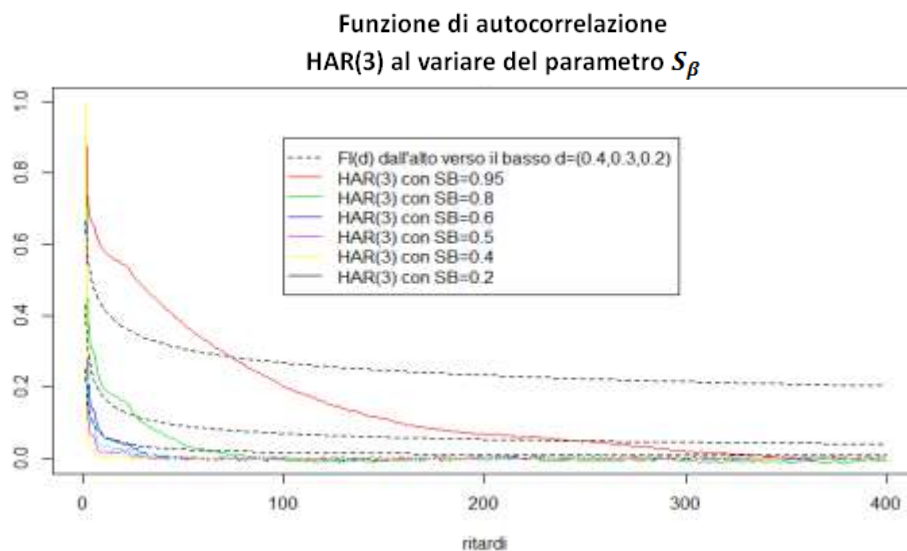


Figura 3 (a)

Funzione di autocorrelazione  
HAR(4) al variare del parametro  $S_{\beta}$

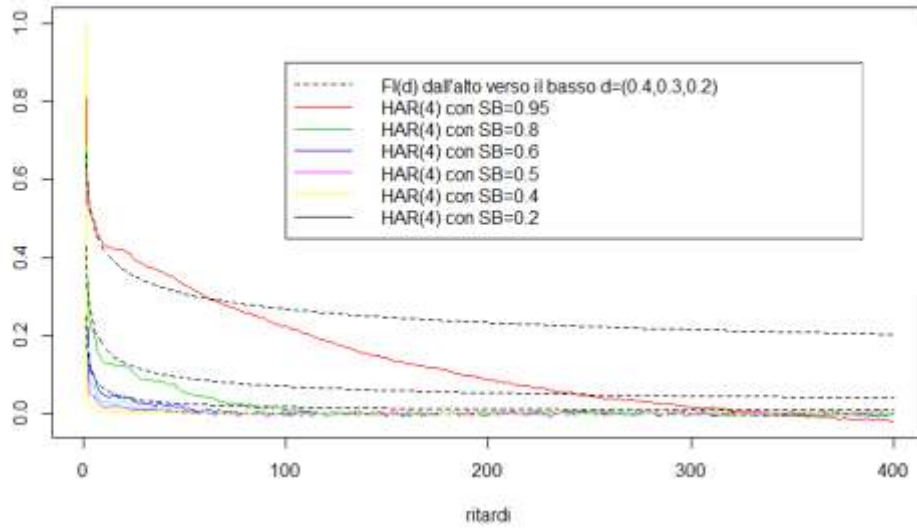


Figura 3 (b)

Funzione di autocorrelazione  
HAR(5) al variare del parametro  $S_{\beta}$

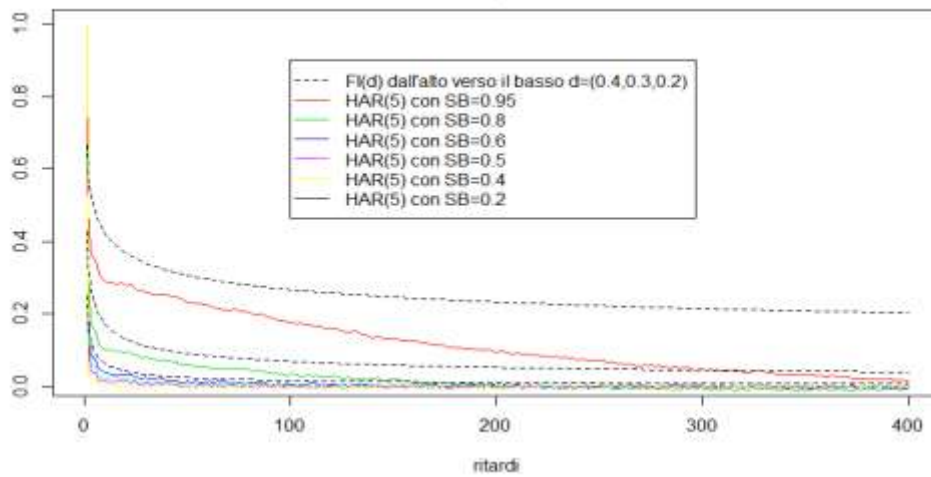


Figura 3 (c)

Funzione di autocorrelazione  
HAR(8) al variare del parametro  $S_\beta$

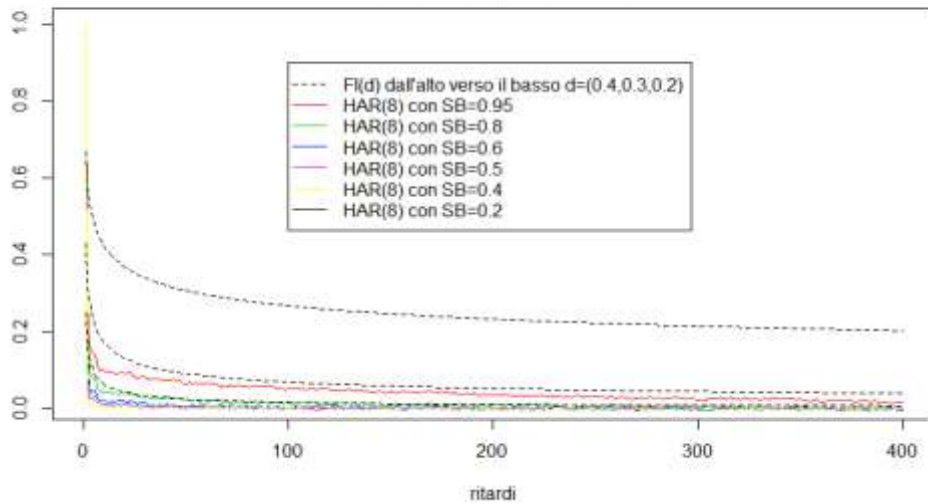


Figura 3 (d)

Come risulta evidente dalla figura 3 (a,b,c,d), per tutti i processi HAR( $p$ ) con  $p=3,4,5,8$ , più la somma dei coefficienti  $\beta_1, \dots, \beta_p$  è prossima a uno, maggiore è la dipendenza di lungo periodo.

Infatti, i  $\beta_j$  esprimono la dipendenza lineare tra  $Y_t$  e  $Y_{t,h_j}$  ed essendo i

$$\beta_j = \frac{S_\beta \lambda^{j-1}}{\sum_{j=1}^p \lambda^{j-1}}, \text{ si nota che minore è } S_\beta, \text{ minore è la memoria del processo.}$$

PARAMENTRO  $\lambda$ :

Per le condizioni necessarie a garantire la proprietà di memoria lunga di un processo HAR( $p$ ) esplicitate tramite (A2), si ha che  $|\lambda|$  deve essere minore di uno.

Inoltre, poiché  $\beta_j \sim c\lambda^j$ , per la natura della variabile risposta che si sta trattando, ovvero la volatilità realizzata,  $\beta_j$  deve essere maggiore di zero con  $j=1, \dots, p$ .

Per questi motivi  $\lambda$  viene fatta variare nell'intervallo (0,1).

Precisamente si sono considerati i seguenti valori di tale parametro: 0.99, 0.85, 0.7, 0.5, 0.3, 0.15,

Si sono presi in considerazione i processi HAR(3), HAR(4), HAR(5), HAR(8).

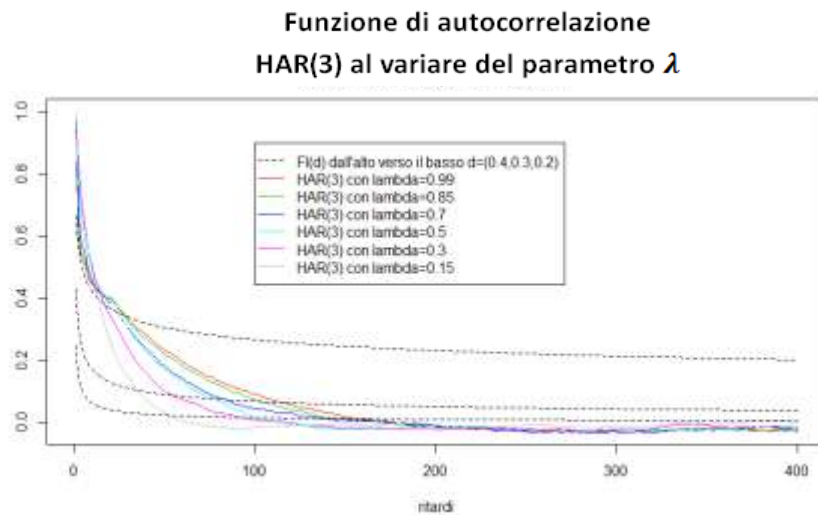


Figura 4 (a)

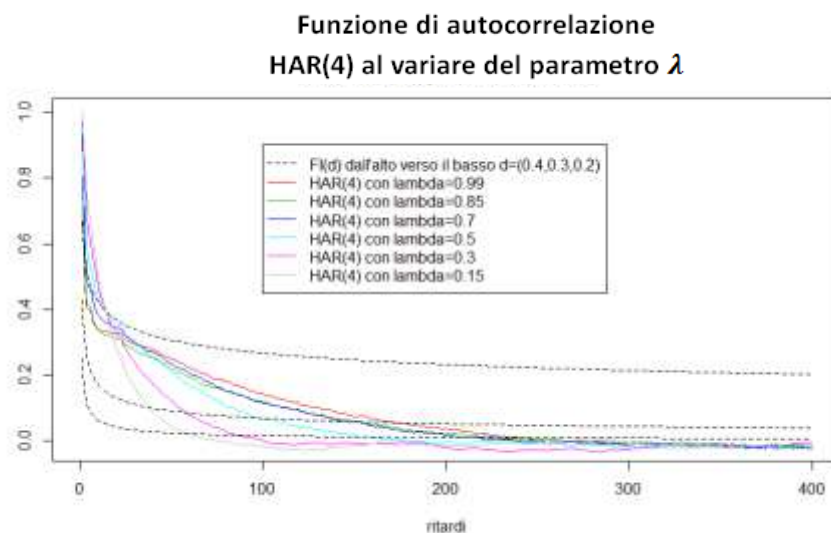


Figura 4 (b)

Funzione di autocorrelazione  
 HAR(5) al variare del parametro  $\lambda$

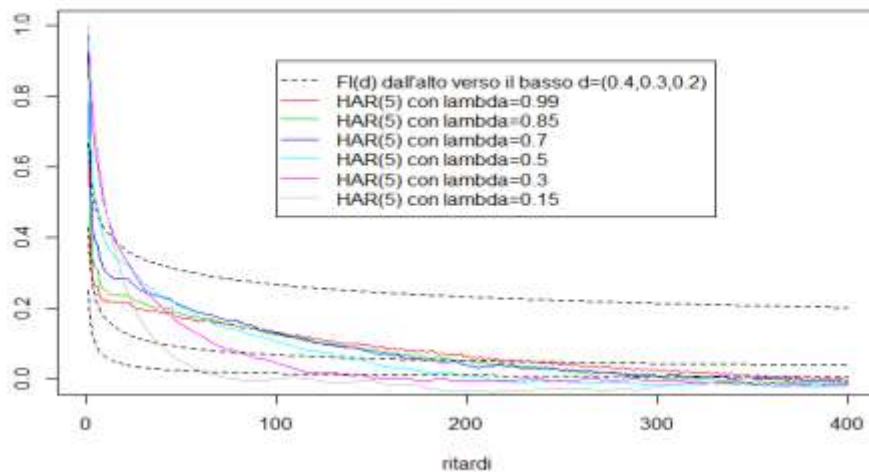


Figura 4 (c)

Funzione di autocorrelazione  
 HAR(8) al variare del parametro  $\lambda$

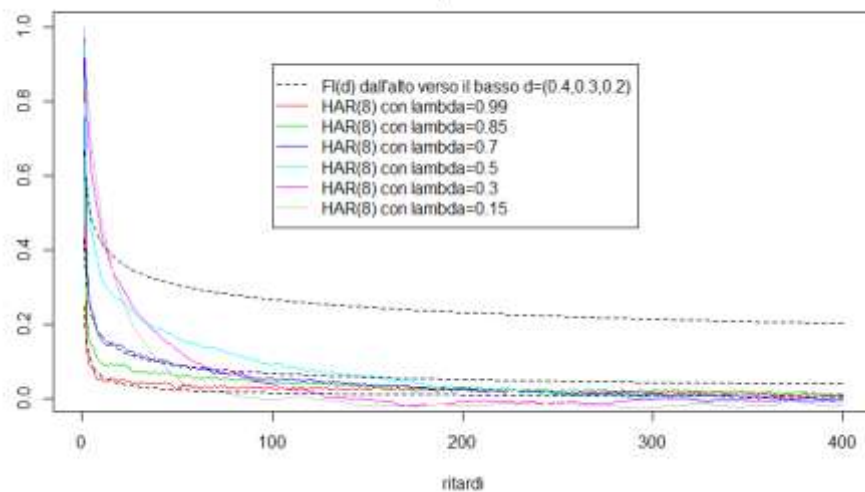


Figura 4 (d)

Dalla figura 4 (a,b,c,d) è interessante notare che le funzioni di autocorrelazione con  $\lambda$  minore, inizialmente sono caratterizzate da una

maggiore autocorrelazione, ma con l'allontanarsi dei ritardi queste ultime tendono a zero più velocemente delle altre.

Per tale motivo si evince che più  $\lambda$  è prossimo a uno, maggiore è la proprietà di memoria lunga del processo. E' importante però notare che questo comportamento diventa meno rilevante con il crescere del valore del parametro  $p$ . Per la suddetta ragione si può ipotizzare che un processo HAR( $p$ ), con  $p > 5$ , possa riflettere la proprietà di memoria lunga anche con valori di lambda prossimi a 0.5.

**PARAMETRO  $\nu$ :**

$\nu$  è quel parametro tramite cui si ricava  $h_j$  con  $j = 4, \dots, p$ , ovvero determina il numero di ritardi della variabile risposta usati nel determinare

$Y_{t,h_4}, \dots, Y_{t,h_p}$ . Ciò significa che tale parametro entra in gioco solo negli HAR( $p$ ) con  $p \geq 4$ .

Per questo motivo a differenza delle precedenti analisi non verrà preso in considerazione il processo HAR(3), ma solo HAR(4), HAR(5) e HAR(8).

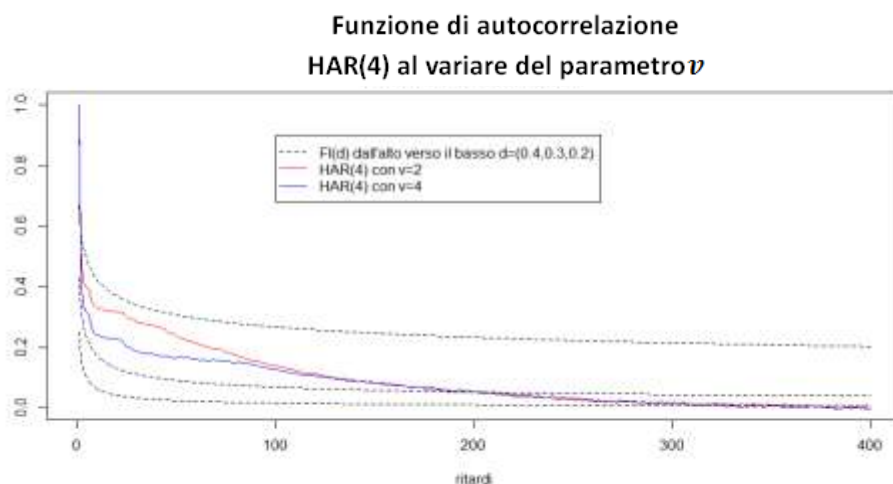


Figura 5 (a)



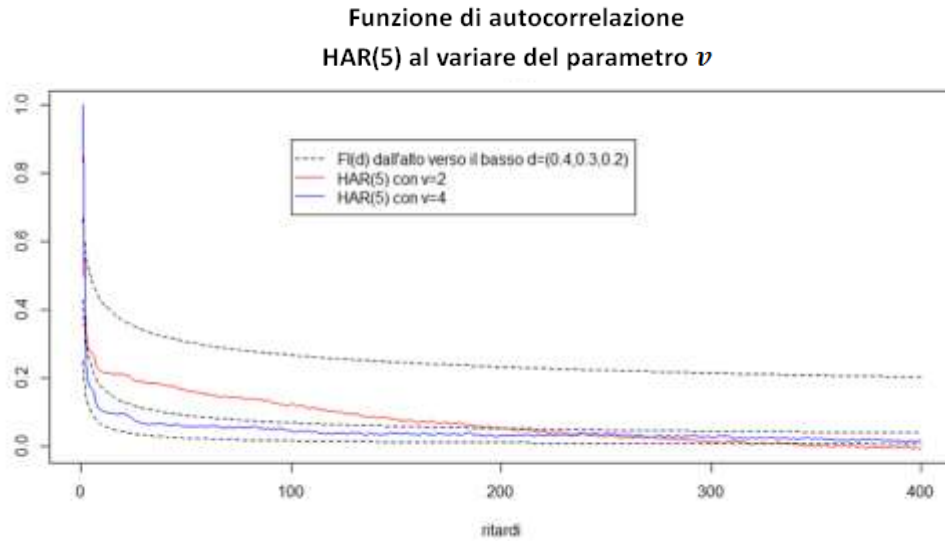


Figura 5 (b)

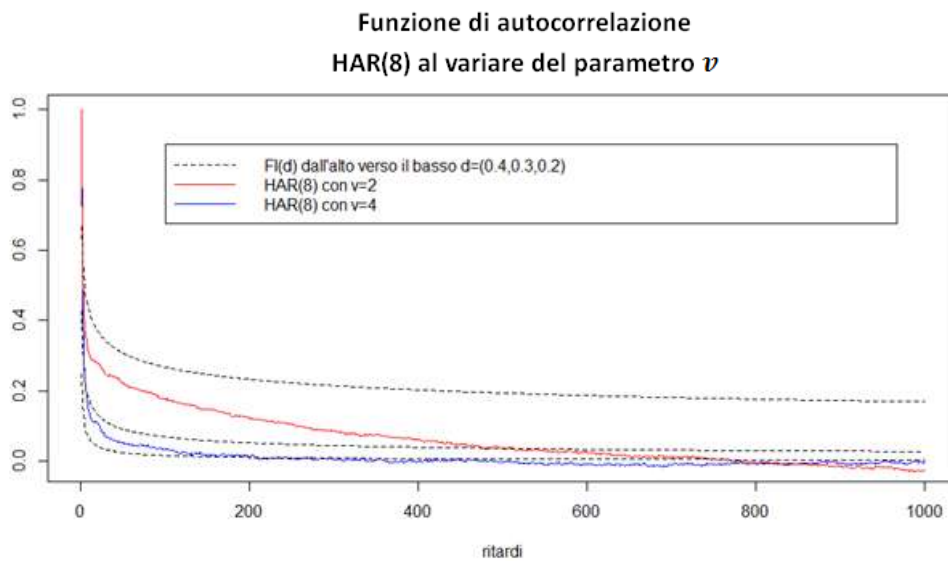


Figura 5 (c)

Dalla figura 5 (a,b,c) è interessante notare che un processo in cui il parametro  $\nu$  è pari a due, sebbene inizialmente presenti una maggiore autocorrelazione, con l'allontanarsi dei ritardi è caratterizzato da una

funzione di autocorrelazione che tende a zero più velocemente rispetto ad un processo in cui  $\nu = 4$ .

Questo significa che più  $\nu$  è elevato, maggiore è la persistenza del processo. E' importante, tuttavia, notare che, all'aumentare di  $p$ , il punto in cui la funzione di autocorrelazione con  $\nu = 4$  incontra quella con  $\nu = 2$ , coincide con ritardi sempre maggiori.

Ne consegue che, per valori di  $p$  elevati, la dipendenza di lungo periodo può essere ben descritta anche da un processo in cui  $\nu$  è pari a due.

# Capitolo 3

## Esperimento Monte Carlo

### 3.1 Obiettivi

In questa sezione si confronteranno, in ambito previsivo, i modelli HAR(3), HAR(4), HAR(5), HAR(8), considerando il caso in cui si disponga di dati generati da un processo a memoria lunga.

L'obiettivo è verificare se, all'aumentare di  $p$ , quindi utilizzando modelli che teoricamente riescono a cogliere maggiormente la dipendenza di lungo periodo, vi sia effettivamente un miglioramento delle previsioni.

Il processo scelto per generare i dati è un HAR(8) la cui specificazione è la seguente:

$$Y_t = \beta_0 + \sum_{j=1}^8 \beta_j Y_{t,h_j} + \epsilon_t, \quad (23)$$

dove:

$$Y_{t,h_j} = \frac{1}{h_j} (Y_{t-1} + Y_{t-2} + \dots + Y_{t-h_j}),$$

$$h_1 = 1, h_2 = 5, h_3 = 22,$$

$$h_j = 22v^{j-3} \quad \text{con } j = 4, 5, \dots, 8, \quad \text{con } v = 2, 4,$$

$$\beta_0 = 0.4,$$

$$\beta_j = \frac{S_\beta \lambda^{j-1}}{\sum_{j=1}^p \lambda^{j-1}} \quad \text{con } j = 4, 5, \dots, 8, \quad S_\beta = \sum_{j=1}^8 \beta_j, \quad \lambda = 0.5, 0.9,$$

$$S_\beta = 0.5, 0.9,$$

$$\epsilon_t \sim N(0,1).$$

Si nota che appositamente si attribuiscono più valori ai parametri  $S_\beta$ ,  $\lambda$ ,  $v$ , perché tramite la loro combinazione si riesce a ottenere un insieme vario di processi a lunga memoria.

Ciò ci permetterà di confrontare la performance delle previsioni dei quattro modelli presi in esame, considerando appunto più processi.

Inoltre, è importante sottolineare che si valuteranno due casi: nel primo si ricaveranno le previsioni partendo da serie di dati, generati tramite la (23), con numerosità pari a 1000, mentre nel secondo si ripeterà l'analisi generando serie storiche con numerosità pari a 10000. Aumentando la numerosità campionaria e disponendo quindi di più osservazioni si cerca di capire se i modelli HAR con ordine più elevato riescano a cogliere questo aumento di informazione producendo così previsioni migliori rispetto al caso in cui la numerosità è pari a 1000.

## 3.2 Procedura di analisi

Per confrontare le *performance* previsionive dei diversi modelli considerati, si utilizza l'errore quadratico medio, definito dalla seguente formula:

$$MSE = \frac{1}{s} \sum_{i=1}^s (Y_{n+\ell}^i - \hat{Y}_{n+\ell}^i)^2,$$

dove  $s$  è il numero di simulazioni,  $Y_{n+\ell}^i$  è il vero valore della  $i$ -esima replicazione al tempo  $(n + \ell)$ , mentre  $\hat{Y}_{n+\ell}^i$  è il valore previsto per la  $i$ -esima replicazione.

Questo indice verrà calcolato per fare previsioni per i seguenti passi in avanti  $\ell = (1, 2, 5, 10, 20, 50, 100)$ .

Per fare ciò, prima di tutto, si sono generate le serie storiche di dati a partire dal modello specificato nell'equazione (23), considerando ogni sua possibile specificazione a cui si perviene combinando i diversi valori dei parametri  $S_\beta, \lambda, v$ .

Le prime osservazioni, necessarie per generare le serie storiche, si sono inizializzate ponendole uguali alla media,  $\mu_y$ , del processo  $Y$ , ovvero

$$\mu_y = \frac{\beta_0}{1 - \sum_{i=1}^p \beta_i}.$$

Si precisa, inoltre, che si sono generate serie storiche la cui numerosità ha la seguente formula:

$$\text{numerosità} = Y_0 + N + NFIT,$$

dove:

$$Y_0 = 100000 ,$$

$N$  = numerosità delle serie, nel primo caso è pari a 1000 mentre nel secondo a 10000 ,

$NFIT = 100$  sono le osservazioni che si prevederanno e tramite cui successivamente si calcoleranno gli MSE agli istanti temporali sopraelencati. Ottenute le seguenti serie storiche, per ognuna si sono scartate le prime  $Y_0$  osservazioni. Ciò viene fatto per assicurarsi che i dati generati non siano dipendenti dalle condizioni iniziali poste, tramite cui si sono inizializzate le prime osservazioni.

Dopo questo procedimento, si ottengono così le nuove serie storiche con numerosità pari a  $N+NFIT$  e usando solo i primi  $N$  dati si sono stimati per ognuna i modelli  $HAR(p)$  con  $p = 3, 4, 5, 8$ .

Attraverso le stime ottenute, si sono potute infine ricavare le previsioni desiderate:  $\hat{Y}_{n+1}, \hat{Y}_{n+2}, \hat{Y}_{n+5}, \hat{Y}_{n+10}, \hat{Y}_{n+20}, \hat{Y}_{n+50}, \hat{Y}_{n+100}$  .

Tali quantità si ottengono per ogni  $HAR(p)$  con  $p=3, 4, 5, 8$ .

Ovviamente, osservando la specificazione della (23), ogni errore quadratico trovato dipende anche da  $\epsilon_t$  che è una realizzazione di una variabile casuale Normale Standard.

Si è ripetuto tale procedimento  $s = 1000$  volte e si sono calcolati così gli errori quadratici medi, tramite la seguente formula:

$$MSE_{n+1}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+1}^i - \hat{Y}_{n+1}^i)^2 ,$$

$$MSE_{n+2}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+2}^i - \hat{Y}_{n+2}^i)^2 ,$$

$$MSE_{n+5}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+5}^i - \hat{Y}_{n+5}^i)^2 ,$$

$$MSE_{n+10}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+10}^i - \hat{Y}_{n+10}^i)^2 ,$$

$$MSE_{n+20}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+20}^i - \hat{Y}_{n+20}^i)^2 ,$$

$$MSE_{n+50}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+50}^i - \hat{Y}_{n+50}^i)^2,$$

$$MSE_{n+100}(p) = \frac{1}{s} \sum_{i=1}^{1000} (Y_{n+100}^i - \hat{Y}_{n+100}^i)^2.$$

Per concludere l'analisi il modello HAR(3) è stato preso come riferimento. Si precisa che la bontà delle previsioni, ottenute dai diversi modelli, si è confrontata mediante il seguente rapporto:

$$\frac{MSE(HAR(p))}{MSE(HAR(3))}, \text{ con } p = 4, 5, 8. \quad (24)$$

Questo risultato evidenzia, se effettivamente l'uso di un modello HAR( $p$ ) con  $p > 3$  abbia fornito previsioni più precise rispetto a quelle ottenute tramite il modello HAR(3). Infatti, per come è definito il rapporto (24), qualora un modello HAR( $p$ ) con  $p = 4, 5, 8$ , produca previsioni migliori di quelle ottenute tramite il modello HAR(3), il valore della (24) sarà minore di uno. Ovviamente si giunge alla conclusione opposta se il rapporto (24) risulta maggiore di uno.

I risultati ottenuti saranno riportati nelle prossime tabelle in cui, si precisa che, per chiarezza espositiva, si sono evidenziati i rapporti con valore minore di uno, al fine di rendere evidenti le situazioni in cui il modello HAR( $p$ ) con  $p = 4, 5, 8$ , preso in esame, produca previsioni migliori del corrispondente modello HAR(3).

### 3.3 I risultati

In tabella 1 sono riportati i risultati nel caso in cui  $N=1000$ :

<b>N=1000</b>									
<b>p=4</b>									
V	LAMBDA	SOMMA	l=1	l=2	l=5	l=10	l=20	l=50	l=100
2	0.5	0.5	0.9501532	1.021697	0.9222583	0.9782997	1.019479	1.080344	0.963201
2	0.5	0.9	1.006126	0.96192	1.060437	1.014934	0.94578	1.01665	1.11771
2	0.9	0.5	1.011416	0.98936	1.039302	1.006595	0.97903	1.01264	1.03566
2	0.9	0.9	1.003913	0.98553	0.956456	0.971823	1.03839	1.01359	1.01619
4	0.5	0.5	1.041655	0.95521	1.044354	1.011766	0.99553	0.9272	1.02336
4	0.5	0.9	1.041465	0.9799	1.035372	0.962592	1.0331	0.9327	0.9905
4	0.9	0.5	1.0362327	1.047591	1.0024098	1.0371830	0.9337440	0.9522154	0.887873
4	0.9	0.9	1.035896	1.052516	1.009187	1.042206	0.9345662	0.9423369	0.890843
<b>p=5</b>									
V	LAMBDA	SOMMA	l=1	l=2	l=5	l=10	l=20	l=50	l=100
2	0.5	0.5	0.958944	1.00561	0.964479	0.981839	0.97127	1.09451	1.03558
2	0.5	0.9	1.045548	0.98722	1.075833	1.101508	0.94936	0.95976	1.08884
2	0.9	0.5	1.054749	0.96243	1.124604	0.991158	0.94457	0.90132	1.07724
2	0.9	0.9	0.960482	0.98495	0.984192	0.92989	1.15952	1.03014	1.03109
4	0.5	0.5	0.958800	0.96971	1.055689	1.035876	0.98489	0.94800	0.95013
4	0.5	0.9	0.958871	0.9791	1.056312	1.050049	0.9985	0.9710	0.9514
4	0.9	0.5	0.997181	1.0043	1.001565	1.007152	1.00778	1.00982	1.00796
4	0.9	0.9	1.047426	0.956417	1.095748	0.9207954	0.9964752	1.091307	1.031241
<b>p=8</b>									
V	LAMBDA	SOMMA	l=1	l=2	l=5	l=10	l=20	l=50	l=100
2	0.5	0.5	0.961726	1.01180	1.04793	1.033552	0.98194	1.00720	1.11491
2	0.5	0.9	1.045233	1.01818	1.170693	1.163137	1.18763	1.24045	1.31327
2	0.9	0.5	1.049186	1.02713	1.117219	1.051801	1.08049	1.04160	1.03301
2	0.9	0.9	0.958599	0.96474	0.963302	0.950515	1.12523	1.13772	1.0952
4	0.5	0.5	0.9730998	0.958223	1.1369568	1.1112989	0.9847647	0.9762059	1.121488
4	0.5	0.9	0.9578521	0.987031	1.0056921	1.1301763	1.0922190	1.1176820	1.175240
4	0.9	0.5	1.0769730	0.948258	1.1208584	1.0355530	1.0332187	0.9990834	1.030524
4	0.9	0.9	1.0215983	0.982653	0.9862152	1.0112563	1.0154695	1.0156931	1.001366

Tabella 1

Osservando i risultati ottenuti per  $p = 4$  si nota che con  $v = 2$ , congiuntamente al variare dei parametri  $S_\beta$  e  $v$ , tutti i valori trovati sono prossimi a uno.



Mentre con  $v = 4$  si nota che i valori inferiori a uno sono situati per lo più in corrispondenza delle previsioni fatte su elevati passi in avanti, nello specifico con  $\ell = (20, 50, 100)$ . Ciò potrebbe significare che il modello HAR(4) riesca a cogliere maggiormente la persistenza dei dati rispetto all'HAR(3). Si evidenzia però, che i valori trovati sono comunque prossimi a uno, quindi la differenza tra le previsioni ottenute con i due modelli potrebbe risultare non significativa. Per questo motivo è opportuno osservare successivamente, se tale comportamento persiste anche nei risultati ottenuti per  $p = 5, 8$ .

Osservando i risultati ottenuti per  $p = 5$  e  $p = 8$  si nota che tutti i valori a cui si è giunti sono prossimi a uno, ad eccezione di alcuni (come ad esempio:  $p = 8, \lambda = 0.5, somma = 0.5, v = 2, \ell = 1$ ;  $p = 8, \lambda = 0.5, somma = 0.5, v = 2, \ell = 50$ ). Anche in questo caso, allora, non sembra che un aumento del parametro  $p$  abbia condotto a previsioni significativamente migliori. Inoltre, osservando i risultati, per tali valori di  $p$  non vi è lo stesso comportamento notato per  $p = 4$ ; infatti i risultati ottenuti per  $\ell = (20, 50, 100)$  sono per lo più maggiori di uno, soprattutto per  $p = 8$ .

Tali esiti ci portano a concludere che non sembrano esservi differenze sostanziali in campo previsivo se si utilizzano modelli HAR( $p$ ) con  $p > 3$ .

In tabella 2 sono riportati i risultati nel caso in cui  $N=10000$ :

N=10000									
p=4									
V	LAMBDA	SOMMA	l=1	l=2	l=5	l=10	l=20	l=50	l=100
2	0.5	0.5	1.005935	0.94502	1.006762	0.927647	0.91037	1.03622	1.08149
2	0.5	0.9	1.006976	0.93709	0.976525	0.918559	0.92657	1.01760	1.0769
2	0.9	0.5	0.974651	1.09443	0.994176	0.885602	1.04492	1.02451	0.92472
2	0.9	0.9	1.211407	0.94861	0.995860	1.026342	0.96152	0.99810	0.97588
4	0.5	0.5	1.053026	1.04238	1.025711	1.060363	0.94105	1.01977	1.06964
4	0.5	0.9	0.876502	0.939033	0.969418	0.981904	0.97094	1.040984	0.92078
4	0.9	0.5	0.986304	0.91658	1.079209	1.083488	1.01398	0.98357	1.02726
4	0.9	0.9	1.044137	1.07032	0.885167	0.901549	1.04762	0.94856	1.00441
p=5									
V	LAMBDA	SOMMA	l=1	l=2	l=5	l=10	l=20	l=50	l=100
2	0.5	0.5	1.053747	1.01512	0.998456	0.998098	0.99710	1.08741	1.1042
2	0.5	0.9	1.053755	0.98923	0.937544	0.969817	0.96894	1.08467	1.11282
2	0.9	0.5	0.956707	1.09664	1.015877	0.995323	1.11198	1.07642	0.96185
2	0.9	0.9	1.142931	0.85931	1.024126	0.924565	0.96660	1.04952	0.94489
4	0.5	0.5	1.10206	1.11810	0.913833	0.962264	0.98497	0.97862	1.07064
4	0.5	0.9	0.965637	0.981794	1.014234	1.018551	1.04880	1.009577	0.93275
4	0.9	0.5	0.869135	0.89210	1.051452	1.094626	0.92946	1.07379	0.96748
4	0.9	0.9	1.027549	0.9893	0.960685	0.972971	1.04805	0.93515	1.03755
p=8									
V	LAMBDA	SOMMA	l=1	l=2	l=5	l=10	l=20	l=50	l=100
2	0.5	0.5	1.001819	1.001172	1.002772	0.997788	0.99834	1.002866	1.00121
2	0.5	0.9	1.007369	0.936912	0.980949	0.912871	0.91999	1.014612	1.07387
2	0.9	0.5	0.97565	1.088556	1.018472	1.040110	1.07761	1.10106	1.04453
2	0.9	0.9	1.195069	0.910410	0.957609	0.998318	1.00852	0.983555	0.96629
4	0.5	0.5	1.055049	1.04553	1.026414	1.061749	0.94359	1.02487	1.07251
4	0.5	0.9	0.966706	0.985062	1.016649	1.020646	1.05197	1.025895	0.9454
4	0.9	0.5	0.975698	0.976611	0.982653	0.971488	0.96524	0.976523	0.95636
4	0.9	0.9	0.956589	0.96214	0.976955	0.965245	0.94859	0.956158	0.91653

Tabella 2

Si nota che i risultati ottenuti in tabella 2 sono anche in questo caso tutti prossimi a uno, ad eccezione di alcuni (come ad esempio:  $p = 4, \lambda = 0.5, somma = 0.9, v = 4, \ell = 1$ ;  $p = 5, \lambda = 0.9, somma = 0.5, v = 4, \ell = 1$ ;  $p = 5, \lambda = 0.9, somma = 0.9, v = 2, \ell = 2$ ;  $p = 4, \lambda = 0.9, somma = 0.9, v = 2, \ell = 1$ ).

Nello specifico per  $p = 4$  e  $p = 5$  non si rileva alcuna sostanziale differenza tra le previsioni ottenute attraverso modelli HAR( $p$ ) e HAR(3).

E' evidente, invece, che per  $p = 8$ ,  $\lambda = 0.9$  e  $somma = 0.5, 0.9$ , sebbene i valori siano prossimi a uno, le previsioni ottenute tramite il modello HAR(8) sono migliori rispetto a quelle a cui si è pervenuti mediante il modello HAR(3) per ogni valore di  $\ell$ .

Questo miglioramento è spiegabile dal fatto che, per la natura dei dati, qualora si usi il modello HAR(8), si ricade nel caso di corretta specificazione.

Per completare l'analisi e osservare se i risultati a cui si è giunti finora vengono confermati o meno, si ripete l'intero studio generando i dati da un processo ARFIMA(0, $d$ ,0), ovvero FI( $d$ ), stazionario, la cui specificazione è la seguente:

$$(1 - L)^d y_t = \epsilon_t , \quad (25)$$

con  $\epsilon_t \sim N(0,1)$  e  $d = 0.15, 0.3, 0.49$ .

Va sottolineato che si considera appositamente il valore  $d = 0.49$ , perché in questo caso si è al limite della zona di stazionarietà (se  $d \geq 0.5$  il processo non è stazionario) e per tale valore si ha un'elevata dipendenza di lungo periodo.

Una volta generati i dati con le varie specificazioni del processo esplicitato nella (25), si trovano, con la stessa procedura applicata precedentemente, gli errori quadratici medi per le previsioni per i seguenti passi in avanti  $\ell = (1, 2, 5, 10, 20, 50, 100)$  relative ai modelli HAR( $p$ ) con  $p = 3, 4, 5, 8$ . Inoltre, si stimano i parametri  $\beta_j$  dei modelli HAR( $p$ ) con  $p = 3, 4, 5, 8$ , i quali differiscono tra loro dalla combinazione dei seguenti parametri

$\nu = 2, 4$ ,  $S_\beta = 0.5, 0.9$  e  $\lambda = 0.5, 0.9$ . Infine, in questo caso, si stima anche il modello FI( $d$ ), per poter così trovare le previsioni e gli errori quadratici medi nel caso di corretta specificazione del modello. Ciò è di fondamentale importanza, perché confrontando gli errori quadratici medi ottenuti tramite FI( $d$ ) con quelli ottenuti prevedendo con un modello HAR( $p$ ), si ricava anche un'informazione aggiunta, ovvero si ha un'indicazione di quanto adatti siano i vari modelli HAR( $p$ ) per prevedere dati che possiedono dipendenza di lungo periodo.

In tabella 3 sono riportati i risultati nel caso N=1000:

<b>N=1000</b>								
<b>P=4</b>								
<b>d</b>	<b>V</b>	<b>l=1</b>	<b>l=2</b>	<b>l=5</b>	<b>l=10</b>	<b>l=20</b>	<b>l=50</b>	<b>l=100</b>
0.15	2	0.922192	0.982739	1.032787	0.961253	0.885986	0.985079	1.05768
0.15	4	0.959659	0.852824	1.042887	1.027415	1.024244	0.991606	1.01628
0.3	2	0.979132	0.986659	1.0958	0.960624	1.050688	1.084395	1.03262
0.3	4	0.962984	1.003564	0.919427	1.027995	0.940577	0.990204	1.04008
0.49	2	0.945835	1.015248	0.936177	1.018417	0.975997	1.053608	0.96745
0.49	4	0.918923	0.869118	1.047592	0.982964	0.970018	1.068383	1.06411
<b>P=5</b>								
<b>d</b>	<b>V</b>	<b>l=1</b>	<b>l=2</b>	<b>l=5</b>	<b>l=10</b>	<b>l=20</b>	<b>l=50</b>	<b>l=100</b>
0.15	2	0.966948	1.036733	0.953080	1.038944	0.908335	1.038113	1.031225
0.15	4	0.907691	0.966356	1.005085	0.988234	0.999795	0.971903	1.002641
0.3	2	0.970823	1.005529	1.031688	0.990995	1.031781	1.139961	1.05472
0.3	4	0.920249	0.932607	0.971137	1.055322	1.028737	1.06888	1.028496
0.49	2	1.083939	1.076263	0.984205	1.155042	1.013388	1.039616	0.925232
0.49	4	0.969714	1.093865	0.970559	1.000325	1.020548	1.030448	1.160393

P=8								
d	V	l=1	l=2	l=5	l=10	l=20	l=50	l=100
0.15	2	0.960598	1.082628	1.013096	1.05942	0.901117	0.967198	1.07608
0.15	4	0.991199	0.925087	1.028656	1.051008	0.994442	1.143852	1.10598
0.3	2	0.918126	1.039953	1.048182	1.109571	1.035526	1.130282	1.09656
0.3	4	0.99901	1.044321	1.024381	1.089057	1.033262	1.058687	1.16194
0.49	2	0.976676	0.948810	0.907007	1.130906	1.095754	1.207605	1.31149
0.49	4	1.003563	0.947898	0.965101	1.084158	1.167255	1.22964	1.35553
FI(d)								
d	V	l=1	l=2	l=5	l=10	l=20	l=50	l=100
0.15	2	0.890677	0.942244	0.976882	0.939692	1.022474	0.95037	1.06517
0.15	4	0.956004	0.974588	0.991113	0.994698	1.002591	1.01256	1.02257
0.3	2	1.092689	1.016175	1.006662	0.950800	0.955841	0.90457	0.92742
0.3	4	1.005665	0.965675	0.895241	1.021087	0.997058	0.92349	1.03449
0.49	2	1.102961	1.21007	1.114397	1.155217	1.162089	1.20359	1.24978
0.49	4	1.020534	1.067719	1.013789	1.015973	1.139492	1.23227	1.38929

Tabella 3

Nella tabella 3 per  $p = 4, 5, 8$ , si confermano i risultati ottenuti nelle tabelle 2 e 1, ovvero anche in questo caso, in cui si generano i dati da un modello  $FI(d)$ , si nota che si ottengono valori prossimi a uno. Quindi le previsioni ottenute tramite  $HAR(3)$  e gli altri modelli  $HAR(p)$  sembrano molto simili.

Anche in questo caso vi sono alcune eccezioni ( ad esempio  $d = 0.15$ ,  $v = 4$ ,  $\ell = 2$ ;  $d = 0.49$ ,  $v = 4$ ,  $\ell = 20$ ;  $d = 0.49$ ,  $v = 4$ ,  $\ell = 50$ ;  $d = 0.49$ ,  $v = 4$ ,  $\ell = 100$ ).

Nell'ultima parte della tabella, dove si confrontano gli MSE ottenuti attraverso il modello  $FI(d)$  con quelli del modello  $HAR(3)$ , si nota che anche in questo caso i valori trovati sono prossimi a uno.

Ciò è rilevante perché, essendo  $FI(d)$  il modello correttamente specificato, tali valori dimostrano che il modello HAR(3) fornisce buone previsioni.

In tabella 4 sono riportati i risultati in cui  $N=10000$ :

<b>N=10000</b>								
<b>P=4</b>								
<b>d</b>	<b>V</b>	<b>l=1</b>	<b>l=2</b>	<b>l=5</b>	<b>l=10</b>	<b>l=20</b>	<b>l=50</b>	<b>l=100</b>
0.15	2	1.036642	1.024081	1.055494	1.018804	1.063775	1.019461	0.9224852
0.15	4	1.187043	1.096295	0.9326739	0.9781851	0.9464601	0.9714022	1.073322
0.3	2	1.011029	0.964937	1.025808	1.025158	0.8579692	1.049616	1.209818
0.3	4	1.021907	0.9580345	1.080041	0.9710194	1.061227	0.9802486	0.988717
0.49	2	1.104749	0.9410712	0.8373335	0.9414384	0.959335	0.9317334	0.936581
0.49	4	0.922765	1.041626	0.9291327	1.062354	1.146922	0.7787253	0.955307
<b>P=5</b>								
<b>d</b>	<b>V</b>	<b>l=1</b>	<b>l=2</b>	<b>l=5</b>	<b>l=10</b>	<b>l=20</b>	<b>l=50</b>	<b>l=100</b>
0.15	2	1.029895	0.9263273	1.021941	0.9866454	0.9448744	1.055399	0.9697592
0.15	4	1.007945	1.047377	0.9363298	0.9855044	0.9984723	0.9503009	1.087819
0.3	2	0.9893332	0.9785198	1.003693	0.9326961	0.8392309	0.9542966	1.180822
0.3	4	0.991932	1.018474	0.8827829	0.9948129	1.094924	0.9677477	0.9759655
0.49	2	1.031723	0.9352461	0.9245703	0.8527434	0.9903666	0.9554767	0.8810962
0.49	4	1.004628	0.962374	0.9114724	1.050796	0.9974484	0.8287525	0.8304686

P=8								
d	V	l=1	l=2	l=5	l=10	l=20	l=50	l=100
0.15	2	1.057958	1.003269	0.9005819	0.9609716	0.9695312	1.170324	1.019464
0.15	4	1.05627	1.090815	0.9846035	0.9607249	0.9279441	1.008387	1.077661
0.3	2	1.017689	1.030377	1.04057	0.999314	0.834577	0.9471182	1.16758
0.3	4	0.951056	1.083256	0.8971929	0.9548783	1.107102	1.092613	1.01216
0.49	2	1.025071	1.0393	0.9812584	0.9286411	1.028879	0.8376109	0.8746769
0.49	4	0.9828423	1.03415	0.931838	1.04945	1.072647	0.8118405	0.8648683
FI(d)								
d	V	l=1	l=2	l=5	l=10	l=20	l=50	l=100
0.15	2	1.053838	1.044211	0.9841715	0.9465192	0.9591529	1.014716	0.9999198
0.15	4	0.9994707	1.082478	0.9747168	0.9876402	1.069227	1.033015	0.9257931
0.3	2	1.085608	1.06171	1.048657	0.9953213	1.070487	1.050816	1.023067
0.3	4	0.9024365	0.9550701	0.9668684	0.9834177	0.9065144	0.9923275	0.8806498
0.49	2	1.044718	1.066853	1.011069	0.9467765	1.027957	0.9287848	0.8844997
0.49	4	1.081289	0.986317	1.052688	0.9472169	0.8454749	1.007745	0.953164

Tabella 4

Anche da questa analisi è evidente che si ottengono risultati con valori prossimi a uno, fatta eccezione per il caso in cui  $p=4$ ,  $d=0.49$ ,  $v=4$ ,  $\ell = 50$ , dove si registra un valore pari a 0.7787253.

Si conferma quindi che le previsioni ottenute mediante il modello HAR( $p$ ) con  $p = 4, 5, 8$ , sono simili a quelle ottenute mediante il modello HAR(3).

Inoltre, è rilevante osservare che anche nell'ultima parte della tabella, nel caso in cui si tratta il modello FI( $d$ ) che è correttamente specificato, i valori sono prossimi a uno.





# Capitolo 4

## Applicazione su dati reali

### 4.1 La serie storica

In questa sezione si prenderà in esame una serie storica reale.

L'obiettivo principale è quello di confrontare i risultati, sia in fase di stima che di previsione, mediante l'uso di modelli  $HAR(p)$  e  $FI(d)$ .

In questa analisi si considera la serie storica dei prezzi dell'indice SP500 con frequenza pari a 15 minuti, osservato dalle ore 8.35 alle ore 15.00 a partire dal giorno 2 gennaio 2004 fino al 31 dicembre 2008.

L'indice S&P 500 è stato realizzato da Standard & Poor's nel 1957 e segue l'andamento di un paniere azionario formato dalle 500 aziende statunitensi a maggiore capitalizzazione.

Fanno parte di questo insieme le azioni di grandi aziende contrattate al New York Stock Exchange (Nyse), all'American Stock Exchange (Amex) e al Nasdaq.

Il peso attribuito a ciascuna azienda è direttamente proporzionale al valore di mercato della stessa.

La serie storica in esame è riportata in figura 6:



Figura 6

Risulta evidente che la serie non è stazionaria in media, per questo motivo si considerano i log-rendimenti.

Per far ciò prima si trasformano le osservazioni in scala logaritmica e successivamente si differenziano mediante la seguente formula:

$$r_{i,t} = \log(P_{i,t}) - \log(P_{i-1,t}), \quad \text{con } t = 1, \dots, N.$$

Si giunge in questo modo alla seguente serie storica:

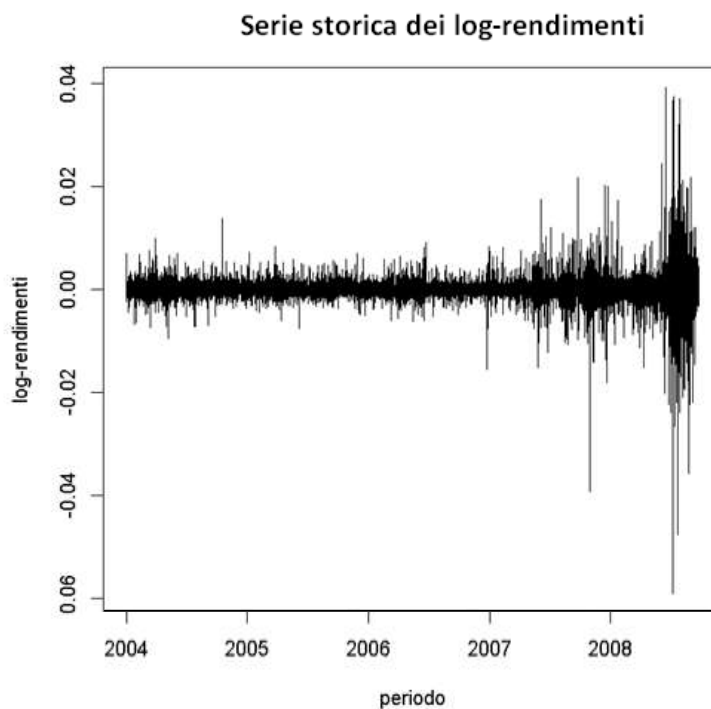


Figura 7

Si nota che la serie storica dei rendimenti, riportata in figura 7, presenta le usuali caratteristiche delle serie storiche finanziarie dei rendimenti ovvero: stazionarietà in media, eteroschedasticità, effetto leva secondo cui shock negativi incrementano la volatilità più di quanto non facciano shock positivi.

Dalla figura 7 è evidente un aumento progressivo della varianza a partire dalla seconda metà dell'anno 2007, che potrebbe essere giustificato a livello economico, dal fatto che gli ultimi mesi del 2007 e in modo più evidente il 2008 sono caratterizzati dal periodo di crisi che tuttora stiamo vivendo. Quest'ultima infatti ha portato ad una maggiore instabilità dei mercati e quindi al progressivo aumento della volatilità dei prezzi e dei rendimenti.

Dopo i risultati iniziali si è ricavata, mediante l'espressione (8), la serie della volatilità realizzata giornaliera, che sarà l'oggetto di studio ed è riportata in figura 8:



Figura 8

Dall'analisi del grafico riportato in figura 8, si confermano i risultati visibili in figura 7, ovvero si nota che considerando il primo periodo fino al 2007, il livello della volatilità realizzata è abbastanza costante e relativamente basso, mentre nel periodo successivo si notano elevate variazioni del valore di quest'ultima.

Dalla Figura 9, in cui si riporta la funzione di autocorrelazione della serie della volatilità realizzata giornaliera, è evidente che la serie storica presenta dipendenza di lungo periodo. Infatti, si osserva che i valori di tale funzione sono significativamente diversi da zero, con un livello di confidenza del 5%, fino a circa 300 ritardi.

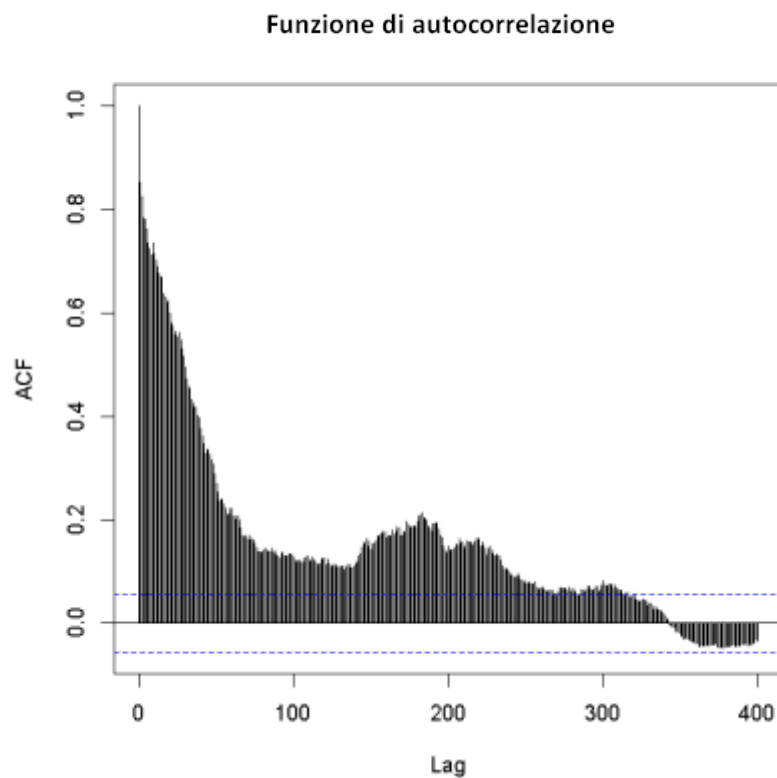


Figura 9

## 4.2 Il confronto tra modelli

Dopo le analisi preliminari si stimano i modelli HAR(3) e HAR(4) la cui specificazione generale è la seguente:

$$Y_t = \beta_0 + \sum_{j=1}^p \beta_j Y_{t,h_j} + \epsilon_t ,$$

dove:  $Y_{t,h_j} = \frac{1}{h_j} ( Y_{t-1} + Y_{t-2} + \dots + Y_{t-h_j} )$ ,

in cui si considerano i seguenti valori:

$$v = 3,$$

$$h_1 = 1, h_2 = 5, h_3 = 22,$$

$$h_j = 22v^{j-3} \text{ con } j = 4.$$

Inoltre si stima anche il modello FI( $d$ ), considerato di riferimento, sapendo appunto che questo è adatto a modellare serie caratterizzate da memoria lunga.

Le stime sono riportate nella tabella 5:

Modello: HAR(2)				
	Stima	s.e.	t-value	p-value
$\beta_0$	0.0005954	0.0001962	3.034	0.00247 **
$\beta_1$	0.3856816	0.0366002	10.538	< 2e-16 ***
$\beta_2$	0.5515342	0.0394709	13.973	< 2e-16 ***
$R^2$	0.7696			
Modello: HAR(3)				
	Stima	s.e.	t-value	p-value
$\beta_0$	0.0004301	0.0002026	2.123	0.03402 *
$\beta_1$	0.3951025	0.0365808	10.801	< 2e-16 ***
$\beta_2$	0.4258141	0.0565895	7.525	1.13e-13 ***
$\beta_3$	0.1358937	0.0439989	3.089	0.00206 **
$R^2$	0.7717			

Modello: HAR(4)				
	Stima	s.e.	t-value	p-value
$\beta_0$	0.0006354	0.0002375	2.676	0.00757 **
$\beta_1$	0.3941294	0.0365555	10.782	< 2e-16 ***
$\beta_2$	0.4035235	0.0581249	6.942	6.72e-12 ***
$\beta_3$	0.2300685	0.0719068	3.200	0.00142 **
$\beta_4$	-0.0996229	0.0601940	-1.655	0.09822 .
$R^2$	0.7736			
Modello: FI(d)				
	Stima	s.e.	t-value	p-value
$d$	0.458	0.0191166	23.95823	< 2.22e-16***

Tabella 5

Per come è costruito il modello, i coefficienti  $\beta_j$  con  $j = 1, 2, 3, 4$ , rappresentano, rispettivamente, la dipendenza lineare tra la variabile risposta e la componente parziale della volatilità giornaliera, settimanale, mensile, trimestrale calcolate con le rispettive osservazioni precedenti al tempo  $t$ .

Analizzando gli output dei seguenti modelli si nota che i coefficienti del modello HAR(2), HAR(3) sono tutti significativi, considerando un livello di significatività del 5%. Inoltre, le stime dei coefficienti sono tutte positive. Ciò è del tutto ragionevole perché, per la natura della variabile risposta e dei regressori usati nel modello, tutti i  $\beta_j$  devono essere maggiori di zero.

A livello interpretativo tale affermazione è facilmente giustificabile dal fatto che un aumento della  $j$ -esima componente di volatilità non può comportare altro che a un aumento del valore della volatilità realizzata giornaliera, quindi la dipendenza tra regressori e variabile risposta deve essere positiva.

Rilevante è l'output relativo all'HAR(4) in cui la stima del coefficiente  $\beta_4$  è negativa, contrariamente alla teoria appena esposta, ma tale risultato non viene considerato perché il coefficiente non è significativo, a un livello del 5%. Ciò significa che la volatilità giornaliera non dipende dalla componente parziale della volatilità trimestrale.

Analizzando nello specifico i coefficienti del modello HAR(2) si nota che  $\hat{\beta}_1 = 0.3856816$  e  $\hat{\beta}_2 = 0.5515342$ ; ciò implica che la variabile risposta è influenzata maggiormente dalla componente della volatilità settimanale rispetto alla volatilità giornaliera calcolata al tempo  $(t-1)$ .

Prendendo successivamente in esame i coefficienti del modello HAR(3), si ha che  $\hat{\beta}_1 = 0.3951025$   $\hat{\beta}_2 = 0.4258141$   $\hat{\beta}_3 = 0.1358937$ .

Risulta così evidente che anche in questo caso  $\hat{\beta}_2$  è il coefficiente più elevato del modello, ne consegue che la variabile risposta ha maggiore dipendenza con la componente di volatilità settimanale rispetto alle altre.

Dato il valore del coefficiente  $\hat{\beta}_1 = 0.3951025$  si nota comunque che anche la dipendenza tra variabile di interesse e il suo ritardo è elevata.

Infine si osserva che  $\hat{\beta}_3$ , pari a 0.1358937, è il coefficiente minore, ma ciò è ragionevole perché esso esprime la dipendenza lineare tra la variabile risposta e la componente mensile di volatilità, perciò è plausibile che questa dipendenza sia inferiore a quella registrata tra variabile risposta e le componenti di volatilità considerate che sono più prossime al tempo  $t$ .



E' rilevante anche evidenziare che tramite l'uso del modello  $FI(d)$  si è ottenuta una stima del parametro  $d$  pari a 0.458. Tale valore, essendo prossimo a 0.5 conferma che la serie storica considerata è caratterizzata da una dipendenza di lungo periodo.

Inoltre, i valori dei coefficienti di determinazione dei modelli HAR stimati, sono rispettivamente 0.7696 , 0.7717, 0.7736.

Tali coefficienti indicano che tutti i modelli stimati sono in grado di spiegare più del 75% della variabilità delle osservazioni, perciò i modelli considerati si possono ritenere buoni.

La figura 10 (a,b,c) riporta le funzioni di autocorrelazione dei residui dei modelli  $HAR(p)$  con  $p = 2, 3, 4$ .

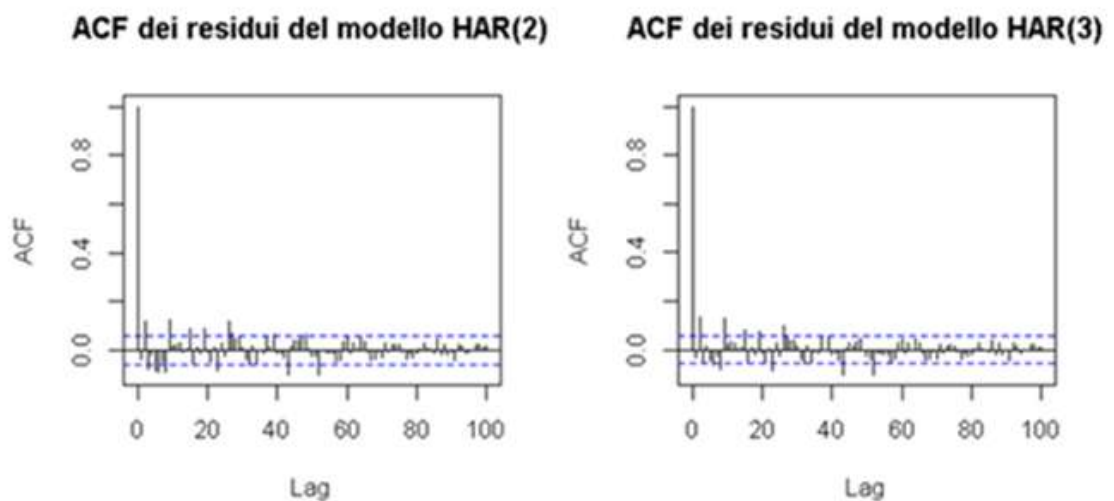


Figura 10 (a)

Figura 10 (b)

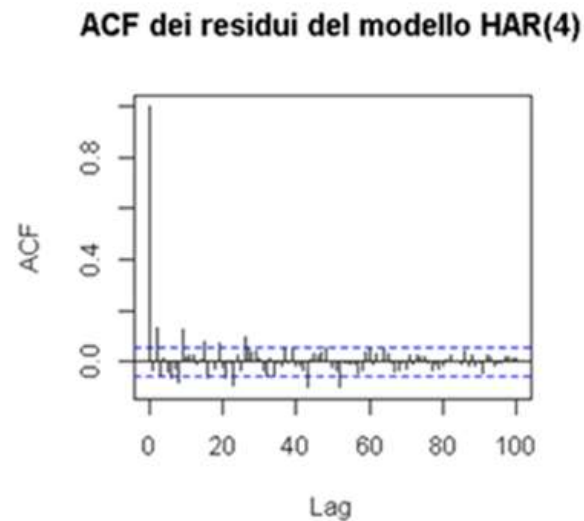


Figura 10 (c)

Risulta evidente dalle funzioni di autocorrelazione, Figura 10 (a,b,c) , che i modelli stimati sono riusciti a cogliere la proprietà di memoria lunga che caratterizza i dati.

Va precisato, comunque, che i residui dei modelli presentano ancora una lieve dipendenza tra la variabile risposta e alcuni suoi ritardi .

Infatti, fino al cinquantesimo ritardo alcuni valori della funzione di autocorrelazione escono dalle bande di confidenza, ricadendo così in casi limite di accettazione o rifiuto dell'ipotesi nulla, considerando un livello di significatività del 5%. Tale dipendenza potrebbe essere colta da qualche modello adatto a modellare l'autocorrelazione di breve periodo, come ad esempio i modelli ARMA.

Continuando il confronto tra i modelli si evidenzia che, poiché il coefficiente  $\hat{\beta}_4$  del modello HAR(4) non risulta significativo, tale modello non sarà preso in considerazione nelle prossime analisi.

Si procede paragonando i modelli mediante il criterio d'informazione di Akaike e il criterio di informazione Bayesiano.

Il modello migliore è quello che presenta valori dei criteri AIC e BIC inferiori.

$p$	AIC	BIC
2	-8595.405	-8575.537
3	-8602.937	-8578.102

Tabella 6

La Tabella 6 riporta i valori di tali criteri. E' evidente che il modello migliore è il modello HAR(3).

Infine, si estende il confronto anche a livello previsivo. A tale scopo si ottengono le stime del modello HAR(2) e HAR(3) usando le prime 1070 osservazioni, che corrispondono all'85% dei dati del campione. In seguito a partire da queste si ricavano le previsioni delle ultime 189 osservazioni.

Tale procedura viene ripetuta utilizzando anche un modello FI( $d$ ). Ciò ha lo scopo di far capire quanto le previsioni ottenute tramite modelli HAR( $p$ ),  $p = 2, 3$  si discostino da quelle ottenute attraverso il modello FI( $d$ ) che, per come è costruito, è adatto a cogliere una memoria di lungo periodo.

Le previsioni dei vari modelli sono state valutate attraverso i seguenti indici:

$$RMSE = \sqrt{\frac{1}{189} \sum_{i=1}^{189} (\hat{y}_{1070+i} - y_{1070+i})^2}$$

$$MAE = \frac{1}{189} \sum_{i=1}^{189} |\hat{y}_{1070+i} - y_{1070+i}|$$

$$MAPE = \frac{1}{189} \sum_{i=1}^{189} \frac{|\hat{y}_{1070+i} - y_{1070+i}|}{y_{1070+i}}$$

I valori di tali indici sono indicati in Figura 11 (a,b,c) e in Tabella 7:

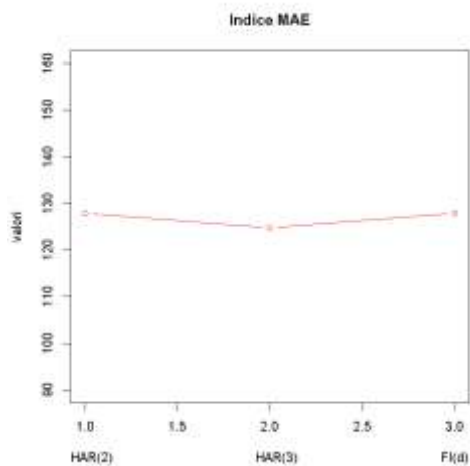


Figura 11 (a)

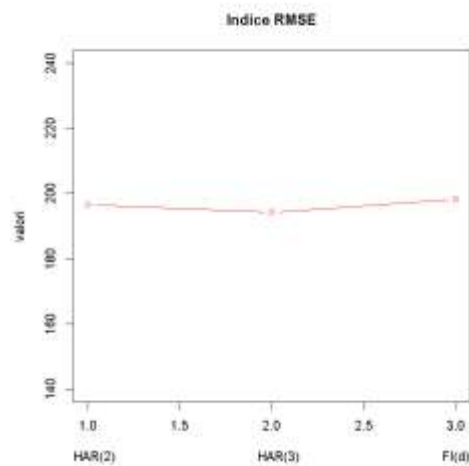


Figura 11 (b)

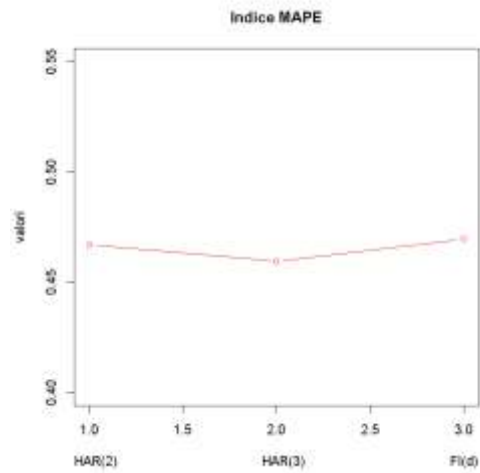


Figura 11 (c)

Modello	MAE	MAPE	RMSE
HAR(2)	127.9318	196.7210	0.4669509
HAR(3)	124.7602	194.2732	0.4594620
FI(d)	127.9254	198.2162	0.4695344

Tabella 7

Dai risultati riportati in Tabella 7 e Figura 11(a,b,c) si evince che il modello HAR(3) risulta il migliore, tra i modelli considerati.

Infatti, si nota che tutti e tre gli indici associati a tale modello hanno un valore inferiore, sebbene di poco, rispetto a quelli calcolati mediante l'utilizzo degli altri due modelli. Inoltre, si evidenzia che secondo l'indice MAE il modello FI( $d$ ) è da preferire rispetto al modello HAR(2), anche se va notato che la differenza è minima, mentre gli indici MAPE e RMSE relativi ai HAR(2) conducono alla conclusione contraria.



# Conclusione

Inizialmente si è esplicitata l'importanza di utilizzare un modello appropriato per trattare dati caratterizzati da una dipendenza di lungo periodo, in modo da riuscire a sfruttare correttamente l'elevato contenuto informativo delle osservazioni.

Dopo aver descritto il processo  $\text{HAR}(\infty)$  proposto in Hwan e Shin (2014), oltre a esplicitare le condizioni di stazionarietà, si sono enunciate anche le condizioni di memoria lunga, secondo cui i coefficienti di tale modello,  $\beta_j$ , decadono esponenzialmente verso zero all'aumentare di  $j$ .

Per tale motivo un processo  $\text{HAR}(p)$ , con  $p$  sufficientemente elevato, può essere una buona approssimazione di un  $\text{HAR}(\infty)$ .

Successivamente, trattando la teoria asintotica dello stimatore dei minimi quadrati, si è dimostrato che lo stimatore OLS è consistente e che asintoticamente si distribuisce come una variabile casuale Normale, risultato su cui si basa l'intera procedura inferenziale.

In seguito, analizzando l'errore di previsione di un  $\text{HAR}(p)$  a un passo in avanti, si evince che in termini previsivi il termine dell'errore che si genera dall'approssimazione di un processo  $\text{HAR}(\infty)$  in un  $\text{HAR}(p)$ , diventa trascurabile.

Analizzando poi la funzione di autocorrelazione di dati generati da processi HAR( $p$ ), si è studiato come le variazioni dei parametri, che costituiscono il processo, influiscano sulle funzioni di autocorrelazione.

Dato il modello specificato nella sezione 2.4 in Corsi (2009) si è giunti a concludere che i processi HAR riflettono maggiormente una dipendenza di lungo periodo al diminuire del parametro  $a$  tale che  $a \in [0,1]$ , mentre considerando la specificazione descritta nella sezione 5.1 in Hwang e Shin (2014), il processo è caratterizzato da una maggiore persistenza all'avvicinarsi dei parametri  $S_\beta \in [0,1]$  e  $\lambda \in [0,1]$  al valore uno, e al crescere di  $\nu$  tale che  $\nu \in N$ .

Osservando i risultati ottenuti mediante la simulazione Monte Carlo si evince che non sembrano esservi differenze sostanziali in campo previsivo tra modelli HAR(3) e HAR( $p$ ) con  $p = 4, 5, 8$ , sia nel caso in cui i dati sono stati generati da un processo HAR(8) che da un processo FI( $d$ ) con  $d$  che assume i seguenti valori: 0.,1 0.3, 0.49.

Inoltre, nel secondo caso, confrontando i risultati del modello FI( $d$ ), correttamente specificato, con quelli del modello HAR(3), si è giunti alla conclusione che quest'ultimo è adatto a modellare la proprietà di dipendenza di lungo periodo.

Infine, anche attraverso l'analisi di una serie storica reale, dove si sono considerati modelli HAR( $p$ ) con  $p = 2, 3, 4$ , si è nuovamente confermato che il modello HAR(3) è il migliore tra quelli considerati.



# Bibliografia

1. Alfarano, S., and T. Lux, 2007. “A noise trader model as a generator of apparent financial power laws and long memory.” *Macroeconomic Dynamics* 11: 80–101.
2. Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001. “The distribution of stock returns volatilities.” *Journal of Financial Economics* 61: 43–76.
3. Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2001. “The distribution of realized exchange rate volatility.” *Journal of the American Statistical Association* 96: 42–55.
4. Arneodo, A., J. Muzy, and D. Sornette, 1998. “Causal cascade in stock market from the ‘infrared’ to the ‘ultraviolet’.” *European Physical Journal, B* 2: 277–282.
5. Barndorff-Nielsen, O., and N. Shephard, 2002a. “Econometric analysis of realized volatility and its use in estimating stochastic volatility models.” *Journal of the Royal Statistical Society, B* 64: 253–280.

6. Barndorff-Nielsen, O., and N. Shephard, 2002b. "Estimating quadratic variation using realized variance." *Journal of Applied Econometrics* 17: 457–477.
7. Corsi, F., 2004. "A simple long memory model of realized volatility." *Manuscript, University of Southern Switzerland*.
8. Corsi, F., 2009. "A simple approximate long memory model of realized volatility." *Journal of Financial Econometrics* 7: 174–196.
9. Gençay, R., and F. Selçuk, 2004. "Asymmetry of information flow between volatilities across time scales." *North American Winter Meetings 90, Econometric Society*.
10. Hosking J.R.M., 1981, "Fractional Differencing."
11. Hwang., E., Shin., D.W., in press, "Infinite-order, long-memory heterogeneous autoregressive model."
12. Ing, C.K., Wei, C.Z., 2003. "On same-realization prediction in an infinite-order autoregressive process." *Journal of Multivariate Analysis* 85: 130–155.
13. Kuersteiner, G.M., 2005. "Automatic inference for infinite order vector autoregressions." *Econometric Theory* 21: 85–115.

14. LeBaron, B., 2006. "Agent-based financial markets: Matching stylized facts with style." *In Post Walrasian Macroeconomics: Beyond the DSGE Model*, ed. D. Colander, 221–235. Cambridge, UK: Cambridge University Press.
15. Lux, T., and M. Marchesi, 1999. "Scaling and criticality in a stochastic multi-agent model of financial market." *Nature* 397: 498–500.
16. Lynch, P., and G. Zumbach, 2003. "Market heterogeneities and the causal structure of volatility." *Quantitative Finance* 3: 320–331.
17. Müller, U., M. Dacorogna, R. Dav, R. Olsen, O. Pictet, and J. von Weizsacker, 1997. "Volatilities of different time resolutions – Analysing the dynamics of market components." *Journal of Empirical Finance* 4: 213–239.
18. Peters, E., 1994. "Fractal Market Analysis." *New York: Wiley*.